1-1-1981

Transition detector circuits for NMOS static memories.

Richard Hogg

TRANSITION DETECTOR CIRCUITS FOR NMOS STATIC MEMORIES

by

Richard Hogg

A Thesis

Presented to the Graduate Committee

of Lehigh University

in Candidacy for the Degree of

Master of Science

in

Electrical Engineering

This thesis is accepted and approved in partial fulfillment of the requirements for the degree of Master of Science.

November 25, 1981
(date)

Professor in Charge

Chairman of the Department

Acknowledgements

I would like to thank Austin C. Dumbri and Walt Rosenzweig for suggesting this topic and for their continuing guidance and encouragement, Dave Beecham and Glen T. Cheney for their support and suggestions, K. H. Lee and the Silicon Processing Technology Laboratory for manufacturing samples, and Bell Laboratories for making this work possible. Finally, my sincerest appreciation to my advisor, Dr. Marvin H. White, for his valuable advice and encouragement.

Table of Contents

List of Tables

List of Figures

Abstract

Address transition activation allows the use of dynamic circuitry in a static NMOS memory, thus reducing power dissipation without increasing access time. This is accomplished without imposing any additional requirements, such as clocking, on users of the device. Circuits to detect address input transitions, called Transition Detectors, must be carefully evaluated to guarantee proper operation of the memory device under all conditions. Two transition detectors are examined and various problems and pitfalls are discussed. It is shown that by careful design, a reliable transition detector may be obtained, and thus the benefits of dynamic circuitry and fully static operation may be simultaneously enjoyed.

1. Introduction

The first semiconductor integrated circuit memory component fabricated with the Metal-Oxide-Semiconductor (MOS) technology which saw widespread commercial use was introduced in 1968 (1). Since then, the MOS technology and minor variations on it have allowed the production of a large assortment of data storage products, including dynamic and static Random Access Memories (RAM's), Read Only Memories (ROM's), and Programmable Read Only Memories (PROM's). These integrated circuits have been constructed using p-channel MOS transistors (PMOS), n-channel MOS transistors (NMOS), and a combination of the two on the same component, in the Complementary MOS (CMOS) technology (2).

Each of these integrated circuits is subject to classification into one of two categories, regardless of whether it is a PMOS, NMOS or CMOS RAM, ROM, or PROM. This classification has to do with the input requirements for operation of the circuit, specifically the requirements for reading the data stored in a certain location on the chip. In some of the circuits, new data may be read merely by changing the address inputs to reflect the address of the desired data. Other components require that the address be changed, and then that an access cycle be initiated by changing the logic level of a special input, called a strobe, clock, or enable input (3).

The latter category of circuits are called synchronous or clocked circuits. In some computer systems the clocked mode of operation is an advantage because it allows the address information to be latched into the memory circuit, under control of the clock input, so that information on the address inputs can disappear before the resulting data is delivered at the output. This mode of operation is incompatible with other computer systems, particularly those using certain microprocessor integrated circuits which are not capable of generating the proper clock signal.

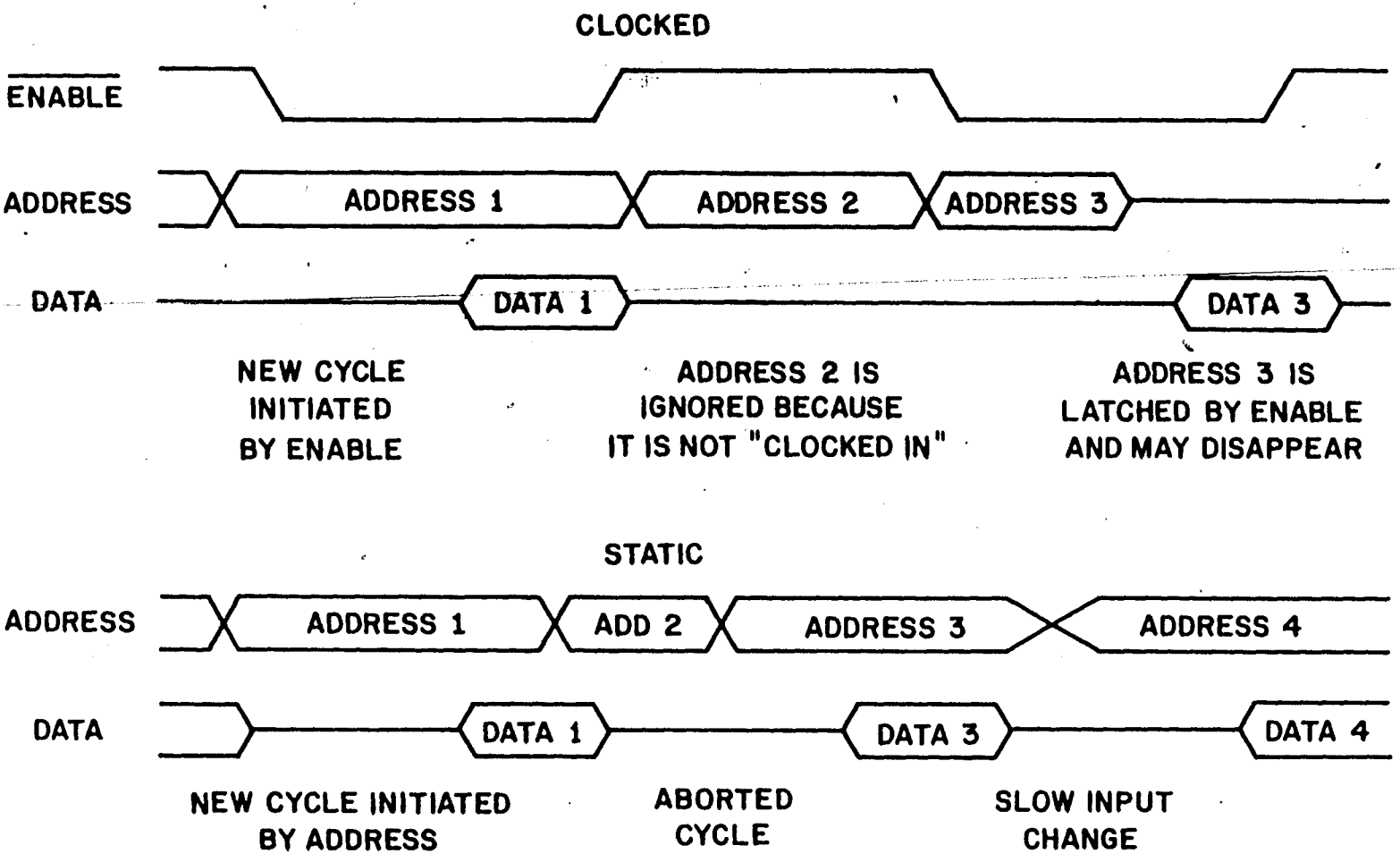
These systems must use components from the other category. These are called asynchronous, unclocked, or static memories. This somewhat unfortunate nomenclature should not be confused with the distinction between static and dynamic RAM's, the former of which has cells which are completely stable as long as power is applied to the circuit, while the latter has cells which must be periodically "refreshed" to prevent the data they contain from fading away. While every dynamic RAM manufactured so far has been clocked, some static RAM's are clocked and some are "truly" static. The clocked static RAM is sometimes said to have a static memory array and dynamic peripheral circuits.

One important difference between a static memory and a clocked memory is that the static memory is guaranteed to operate provided only that a minimum of timing requirements are satisfied on its in-

puts, while a clocked memory usually requires that a much more stringent set of requirements be satisfied. These differences are illustrated in Figure 1. In particular, limits are often placed on the rate at which the inputs of a clocked component must change - for example, the inputs of a particular clocked memory might be required to change from one state to another in no more than thirty nanoseconds. Such limitations are usually not placed on static memories. Also, clocked memories often have a minimum cycle time specification, that is a minimum time which one must wait after initiating an access cycle before initiating another access cycle. This requirements does not allow an access cycle to be aborted. In a static memory, aborted cycles are allowed. Regardless of what has happened before the beginning of an access cycle, the component is guaranteed to deliver the correct information a certain time after the desired address has become stable on the address inputs.

While the requirements of clocked operation may be troublesome to the designer of a computer system, they are an advantage to the memory integrated circuit designer. They allow the use of advanced circuits which operate faster and consume less power than would otherwise be possible.

A design technique has been developed which allows the use of advanced clocked circuits in a component which operates like a fully static memory. This technique uses a circuit called a Transition



OPERATIONAL DIFFERENCES BETWEEN CLOCKED AND STATIC MEMORIES
FIGURE 1

Detector, which is the subject of this work. A memory using this technique is called an Address-Transition-Activated Memory.

The remainder of this introduction provides a brief comparison of the performance of various static RAM's to date. Following this, Chapter Two details the design of a static memory and the means of improving its performance using transition detectors. Chapter Three presents the design of two transition detector circuits and explains the Computer-Aided Design (CAD) techniques used in the development of these circuits. Chapter Four provides information on the fabrication of these circuits and some experimental results. Finally, Chapter Five includes a summary and some conclusions.

Comparison of Static RAM performance. We investigate results on Static RAM's because they represent the bulk of the interesting work on static memory design in the MOS technology to date. Clocked static RAM's and those using unconventional technology are excluded from this comparison for clarity and to focus on the topic of this work.

Also excluded are memories larger than 1,024 bits which have only one data input/output bit. This is because most static RAM's of 16,384 or more bits are organized as 2,048 or more 8-bit words, for convenience of use in small computer systems. However, memories organized in 1-bit words have an advantage in their apparent performance, primarily because they do not require as many data output drivers, which use up considerable power.

Size	Technology	Minimum Design Rule microns	Transition Detectors?	Access Time	Power	Power×Delay / Bit	Ref.
1KX4	NMOS	3.5	no	135nS	220mW	7.25pJ	4
2KX8	NMOS	3.0	no	150nS	600mW	5.49pJ	5
1KX1	NMOS	3.0	no	15nS	330mW	4.83pJ	6
1KX1	NMOS	3.0	no	15nS	320mW	4.69pJ	7
2KX8	NMOS	2.8	yes	100nS	335mW	2.04pJ	8
4KX8	NMOS	3.0	yes	150nS	350mW	1.60pJ	9
2KX8	CMOS	3.0	no	95nS	200mW	1.16pJ	10
2KX8	CMOS	4.0	no	74nS	200mW	.90pJ	11
2KX8	CMOS	3.0	no	74nS	190mW	.86pJ	12
2KX8	CMOS	5.0	yes	150nS	14mW*	.85pJ	13
8KX8	CMOS	2.0	no	80nS	300mW	.36pJ	14

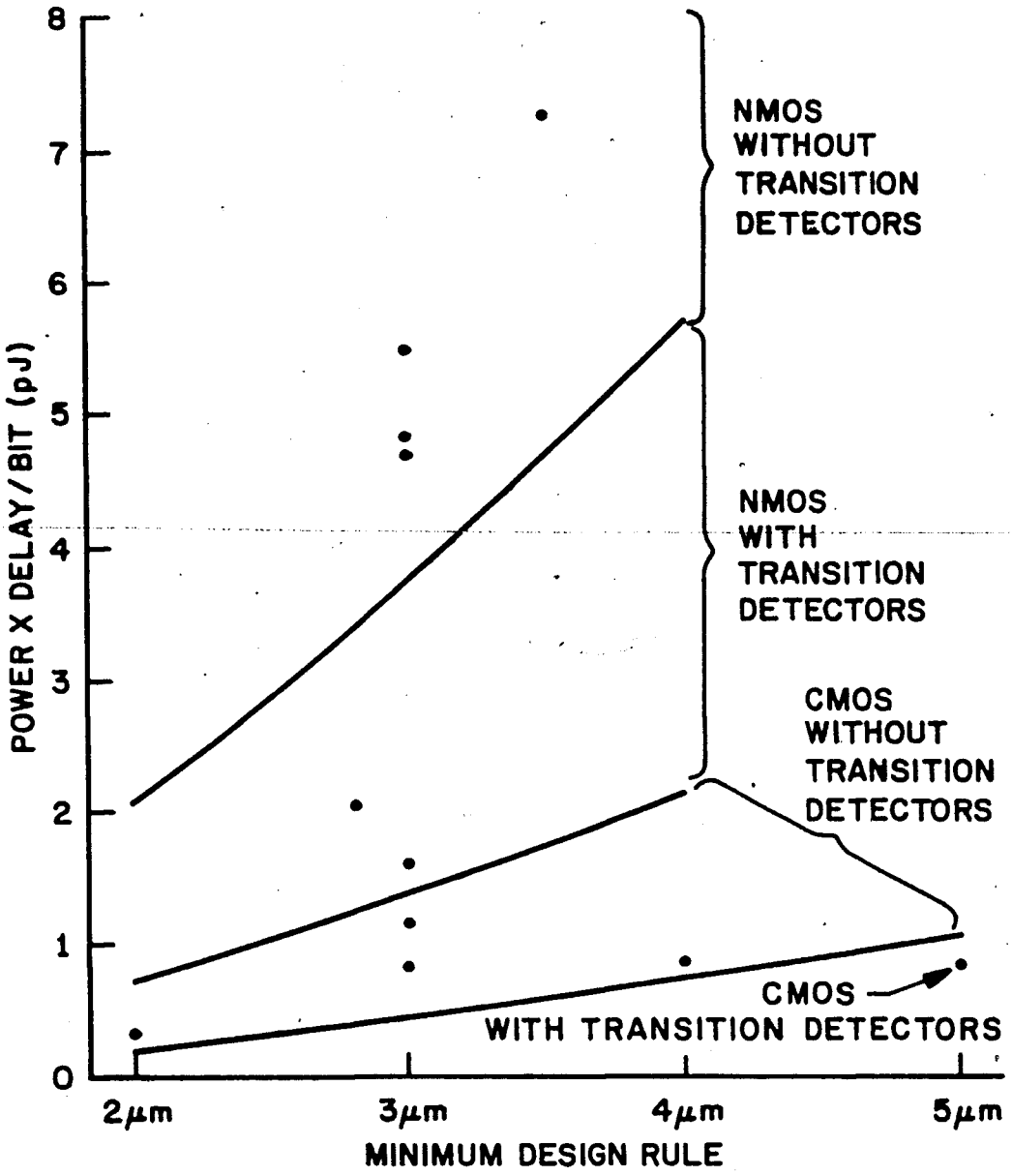
*Specified for 1MHz cycle, therefore power delay/bit=14mW×1μs/16384

Table 1. Comparison of previous SRAM designs

Table 1 lists some important characteristics of eleven static RAM's which are subject to this comparison. included are 1,024 word by 1 bit, 1,024 word by 4 bit, 2,048 word by 8 bit, 4,096 word by 8 bit, and 8,192 word by 8 bit SRAM's. Listed for each memory are the access time, the power dissipation, the technology, the minimum design rule, and whether or not the design uses transition detectors. The access time is the time the memory typically requires to supply output data after it receives new address information. The power dissipation is the average amount of power the component uses, running at maximum speed unless noted otherwise. The minimum design rule is a measure of the photolithographic resolution available in manufacturing the component.

Also included in the table is a figure of merit, the product of the access time and the power dissipation divided by the number of bits. This figure of merit is a general indication of how good the design is in terms of high speed and low power. A low power-delay product per bit is desirable.

Figure 2 shows each of these memories on a graph of power-delay product per bit versus minimum design rule. This graph clearly indicates the advantage of "tighter" design rules in achieving a better performance with a given technology and design technique. It also shows the advantage of CMOS technology over NMOS. However, for a given minimum design rule in a given technology, the graph also clearly indicates the advantage of using transition detectors.



PERFORMANCE COMPARISON OF STATIC RAM'S

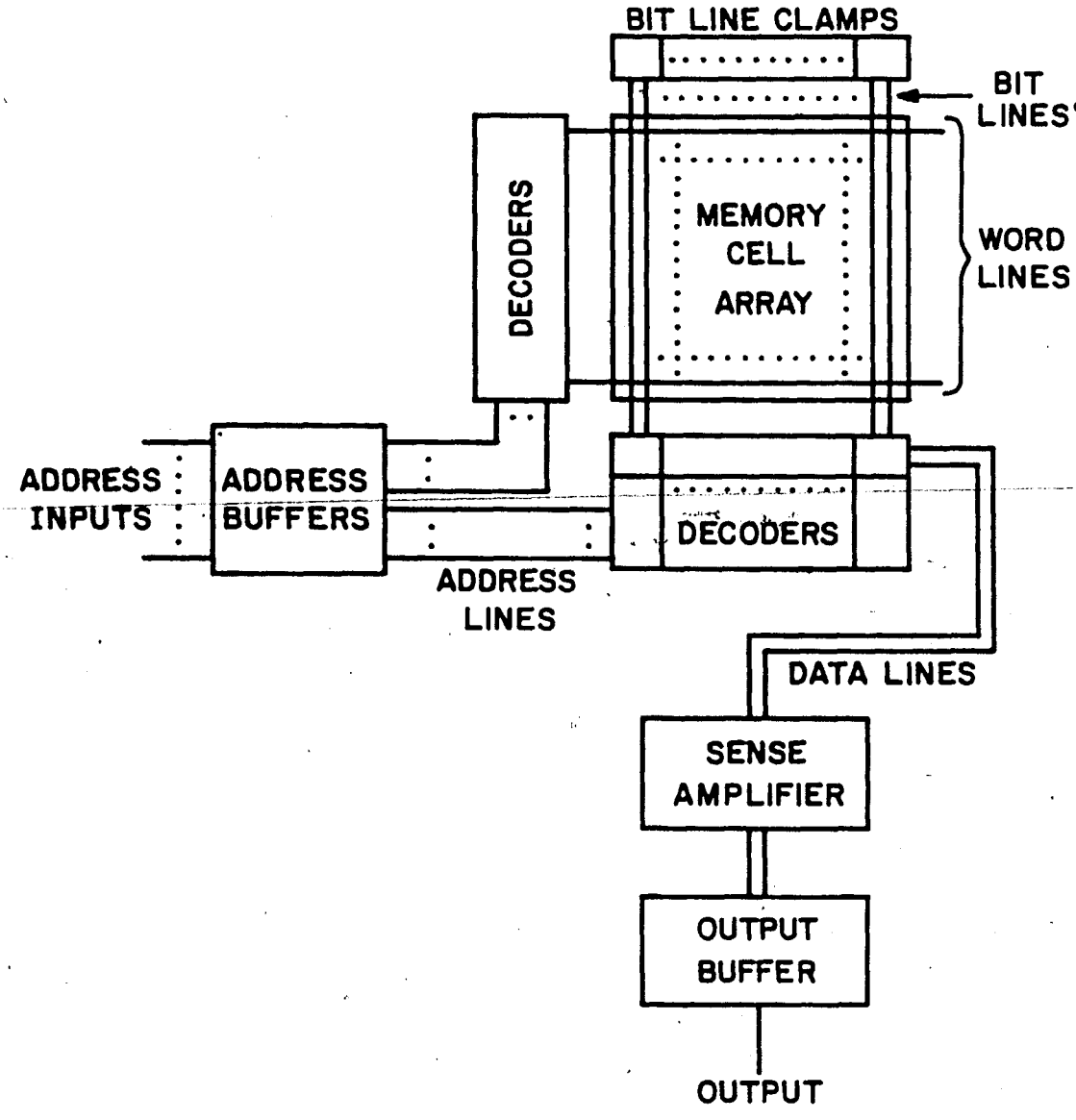
FIGURE 2

2. The Design of an NMOS Static Memory

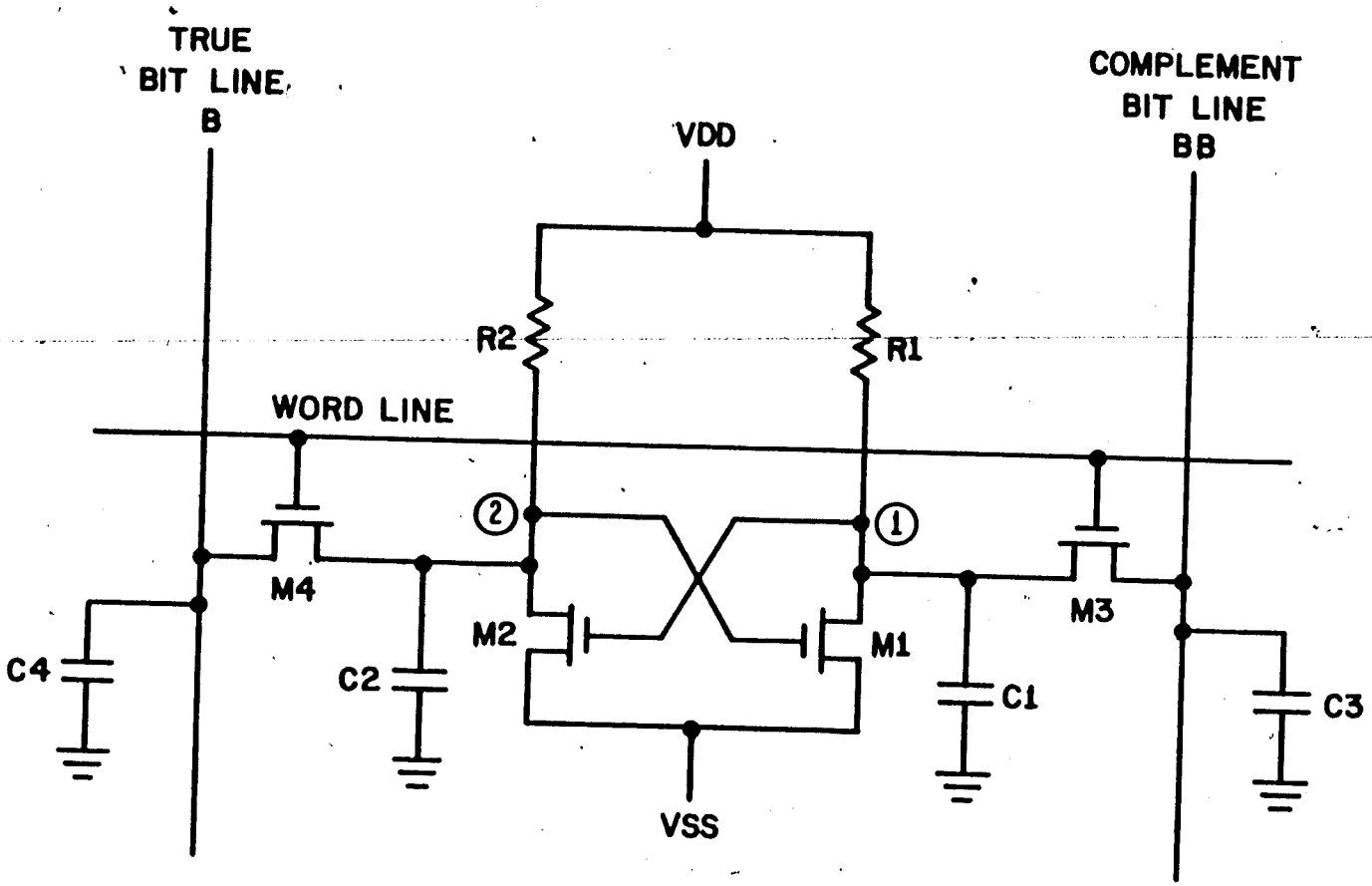
It has been shown that static memory designs which used transition detectors have achieved better performance than those which did not. To examine the reasons for this, this chapter will investigate the design of an NMOS static memory. First the circuitry in a typical static RAM which does not use transition detectors is presented. Then several of the peripheral circuits in this RAM are examined and lower power clocked circuits are presented as alternatives. Finally the address-transition-activation scheme is explained as a means of incorporating the clocked circuits into the static RAM design.

The Fully Static RAM. A block diagram of a fully static NMOS RAM is shown in Figure 3. Certain portions of the memory are omitted from this diagram. These are the circuits used in writing the memory, that is, those circuits used in placing new data into the memory cells. Details of this circuitry are not necessary for our discussion.

A schematic diagram of the memory cell is shown in Figure 4. Vss and Vdd are the negative and positive power supply terminals to the integrated circuit. Vss is considered to be ground. The memory cell consists of cross-coupled enhancement mode MOS transistors M1 and M2, pullup resistors R1 and R2, and access transistors M3 and M4. The cell has two stable states, one of which is used to store a one and the other is used to store a zero.



STATIC RAM BLOCK DIAGRAM
FIGURE 3



STATIC RAM CELL

FIGURE 4

If the cell contains a zero, node 1 is at a potential near Vdd. Since node 1 includes the gate of M2, it is turned on. When on, M1 or M2 has a much lower resistance than R1 or R2. Thus when M2 is on, the potential of node 2 is near Vss.

Now, since node 2 includes the gate of M1, it is turned off. This allows node 1 to remain close to Vdd, and thus the cell is in a stable state.

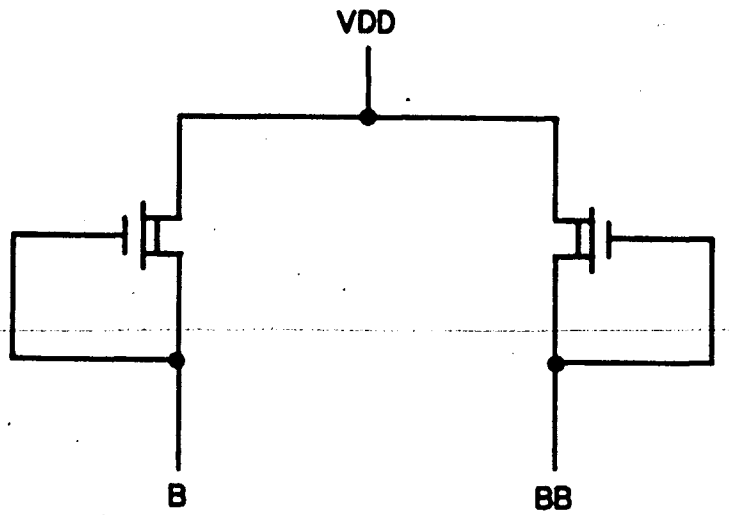
For storage of a one, node 1 has a potential near Vss and node 2 has a potential near Vdd. This state is stable too.

This discussion assumes that the potential of the row line or word line is near Vss, and thus that the access transistors are turned off. This is the proper condition of the word line for simply storing data. To read out the data, the potential of the word line is brought to near Vdd, allowing the cell to impress its data onto the column or bit lines.

It is important to note two facts at this point. First, the bit lines are common to all of the cells in any given column of the memory array; thus the bit lines have a large capacitance to various other nodes compared to the capacitance of node 1 or node 2. These "parasitic" capacitances are indicated by C1, C2, C3, and C4. Second, the resistance of R1 and R2 are very large. This is so that the power dissipation of the memory array can be kept as low as possible.

What happens when the word line potential is raised, if the potential of one or both of the bit lines is near Vss? Suppose the cell stores a zero, and the potential of node 1 is near Vdd. Now suppose the potential of the complement bit line (bit bar, or BB) is near Vss when the word line potential is raised. R1 cannot supply the current to charge C3 to Vdd as the word line is raised, and thus node 1 is pulled to Vss potential by the complement bit line, and the information on the cell is lost. On the other hand, if both bit lines are at Vdd potential, M1 or M2, whichever is turned on can supply more current to discharge C3 or C4 to Vss. It is usually arranged that M1 and M2 are wider, higher gain devices than M3 and M4 so that the low node of the cell, that is whichever of node 1 or node 2 is near Vss, cannot be pulled up as the word line potential is raised. Thus it is important that the potential of both bit lines be high (near Vdd) before the cell is accessed.

The Bit Line Clamps. It is the function of the bit line clamps to raise the bit line potential to near Vdd between the time when one cell is accessed and the time when the next cell is accessed. In the purely static RAM the bit line clamp is simply a load resistor connected between the bit line and Vdd. This load resistor may be nonlinear and is often conveniently implemented by a depletion-mode MOS transistor (that is, a transistor with a negative threshold) as shown in Figure 5. Note that the symbol for the depletion tran-



STATIC BIT LINE CLAMPS

FIGURE 5

sistor differs from the symbol for the enhancement transistor in that the depletion transistor has an extra line from source to drain.

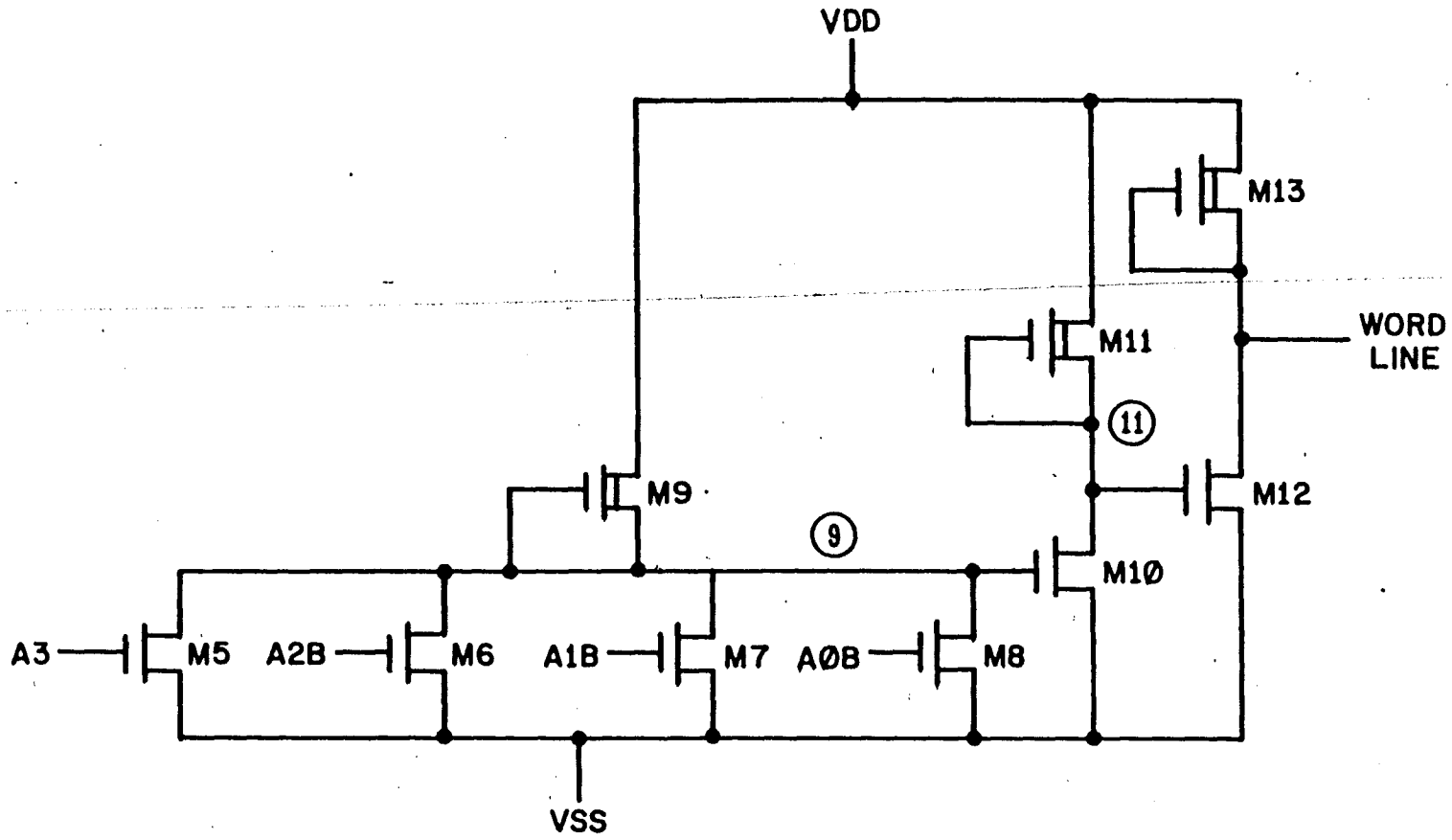
Obviously the conductance of the bit line clamp must be lower than the conductance of the cell access transistor, so that the bit line can be pulled down. However, the conductance of the bit line clamp must be sufficient so that the bit line potential can be recovered in the short time allowed between between accessing one cell and the next. Now consider the situation during the time when a cell is accessed: one of each pair of bit lines is pulled to some potential below Vdd, and thus current is flowing through half of the bit line clamps. A 16,384-bit RAM is typically organized in 128 rows and 128 columns, so that current is flowing through 128 bit line clamps. This represents significant power dissipation.

It would be nice to have a bit line clamp which turned on during the time between cell accesses to recover the bit lines to Vdd, but then turned off during the time a cell is accessed to save power. This would require a signal to tell the clamps when to turn on and off. Such a signal may be derived from a transition detector, as will be described later.

The Decoders and Address Buffers. For a given address input into the memory, exactly one word line must be raised and exactly one pair of bit lines (or, in the case of an n-bit word memory, one set of n pairs of bit lines) must be coupled to the outputs. It is the function of the decoders to perform this selection. A static decoder is shown in Figure 6. This decoder raises its word line when Address 3 (A3) equals 0 and Addresses 2, 1, and 0 equal 1. When this occurs the complement of Addresses 2, 1, and 0 (A2B, A1B, and A0B) equal 0. When this occurs M5, M6, M7, and M8 are all off. Node 9 is thus pulled to Vdd by M9, turning on M10 and pulling node 11 to Vss. M12 is thus turned off and the word line is pulled up by M13.

M13 must be a fairly wide, conductive device to pull up the word line in the time required. It is for this reason that the word line is not connected directly to M5, M6, M7 and M8, because then they would have to be quite large to pull down the word line against M13. This would result in a very large capacitance on the address lines, and since this capacitance is charged and discharged every time the address changes, more power dissipation would result. Instead, small devices are connected to the address lines and node 9, and M10, M11, M12 and M13 are used as a buffer.

M10 and M11 form a common circuit called an inverter. This circuit outputs the logical opposite of its input. If node 9 is



STATIC DECODER

FIGURE 6

low, M10 will be off and M11 will pull node 11 high. On the other hand, if node 9 is high, M10 turns on. It is sufficiently large enough to overcome the pullup action of M11 and pull node 11 low. In the region of the transition between high and low, this circuit has both voltage gain and power gain. The connection of inverter M10 and M11 to inverter M12 and M13 forms a noninverting amplifier or buffer.

Just as exactly one word line is pulled to Vdd, all the rest (127 in a typical 16,384-bit memory) are held at Vss. Thus 127 "M13's" are dissipating power with their source pulled to Vss. This is another area where significant power can be saved by using transition detectors, as will be explained later.

The address buffers have the function of generating the complement addresses from the address inputs. They also amplify the input address so that the decoders are presented with a signal very close to Vss or Vdd, whereas the input address need only conform to the Transistor-Transistor Logic (TTL) logic level specification (15) of: logic 0 less than 0.8 volts, logic 1 greater than 2.4 volts.

The Sense Amplifier. The sense amplifier is a differential amplifier which converts a small differential signal on the bit lines to a large signal on the data lines. (The selected pair of bit lines are coupled to the data lines by a decoder.) Without the sense amplifier, the access time would have to include time for the

very small cell transistor to pull the very large bit line capacitance below one threshold voltage. A memory which took this long to access would not be useful to most systems.

In order to amplify the small differential signal on the bit lines to full power supply levels at the required speed, the sense amplifier must contain some fairly large devices biassed into their linear region. These devices are usually organized as several stages of differential pairs. These gain stages need to dissipate significant power to be effective. Details of the circuitry are not necessary for this discussion. An address-transition-activated memory can use a simple five-transistor sense amplifier, described later, which consumes much less power.

The Output Buffer. The output buffer has the job of driving the output data onto the external lead and whatever capacitance is connected to it. Most memory components are specified to drive 50 picofarads of capacitance on the output. In addition, the output buffer must be able to drive a TTL input, and thus must be capable of sinking 1.6 milliamperes to Vss.

While a schematic diagram of an output buffer is not presented here, it is easy to understand how the output buffer requires power. It must contain two very wide devices, one to pull up the output lead to output a one, and one to pull down the output lead to output a zero. The gate of each of these devices must be connected to some

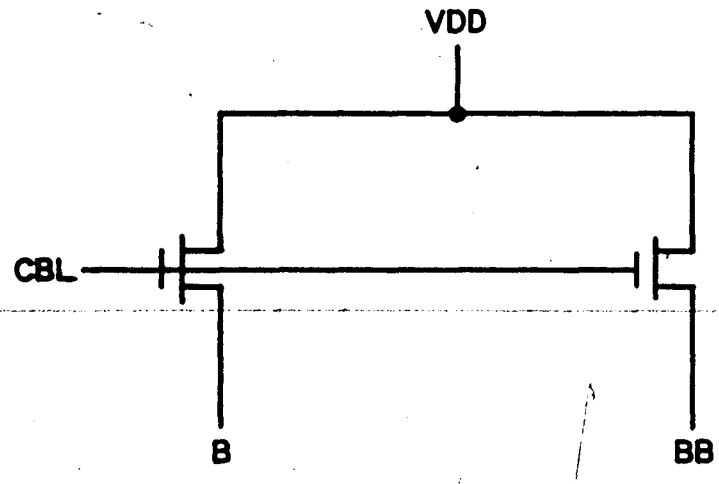
sort of pullup resistor and pulldown transistor capable of moving the gate to its proper potential quickly. Obviously at least one of the two gates must be near Vss at all times. Here, as in the case of the decoder, the pullup resistor connected to the turned-off gate is dissipating power.

Dynamic Circuit Alternatives. This section presents alternatives to several of the circuits described in the last several sections. Each of these circuits depends on the presence of a clock signal, the generation of which will be described in the next section.

The dynamic bit line clamps are shown in Figure 7. They are simply source-follower transistors which pull the bit lines to one threshold voltage less than CBL (Clock Bit Lines), unless the bit line is at a higher potential than that. CBL must be at a high potential during the time between accessing one cell and the next, and it must be at a low potential when a cell is accessed. It is sufficient that the bit lines are recovered to one threshold below Vdd instead of all the way to Vdd.

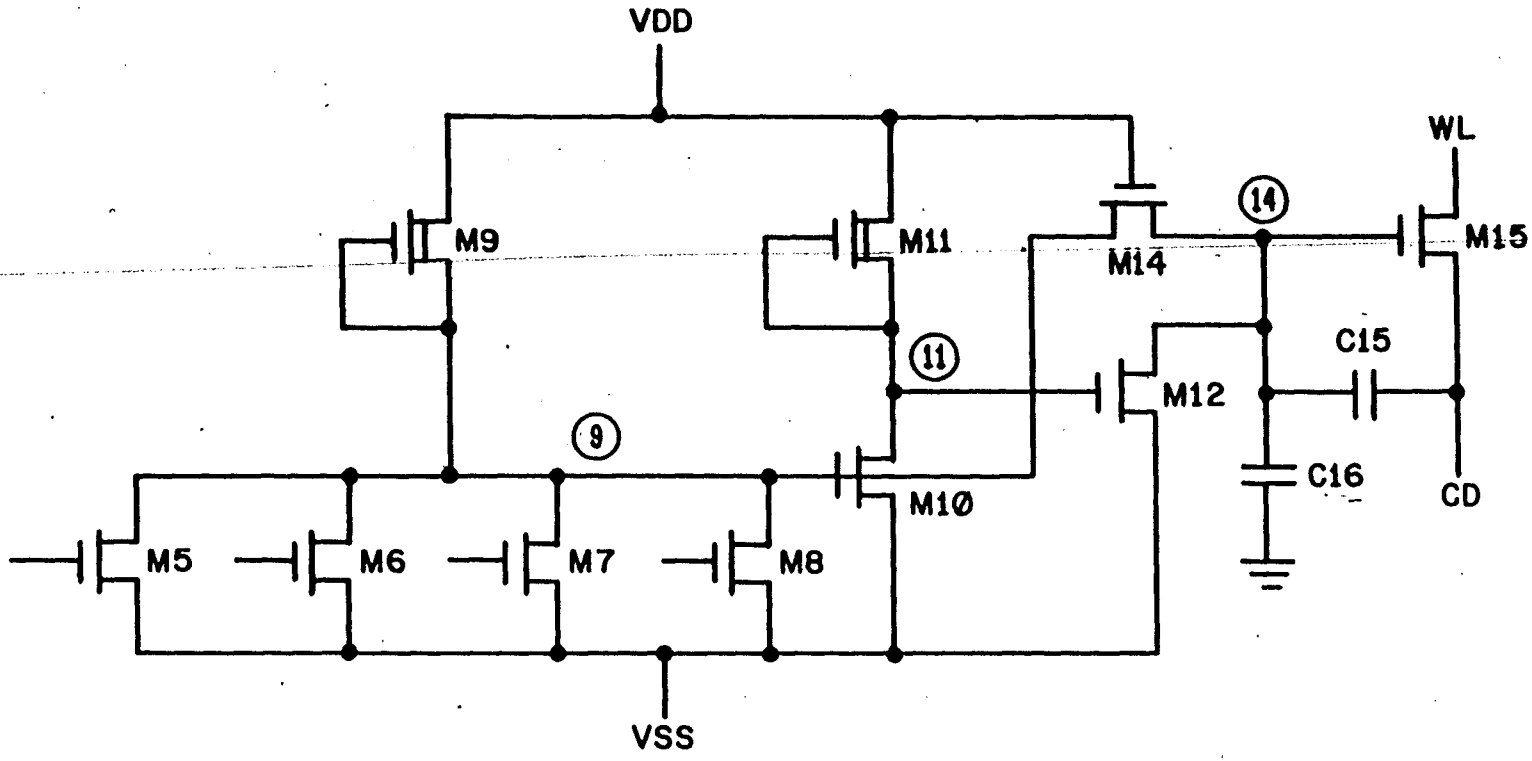
A dynamic decoder is shown in Figure 8. In this circuit M13 is replaced by M14 and M15, forming what is known as a bootstrap circuit. Operation of this circuit is as follows.

When the address lines are changing to new values and shortly thereafter, CD (Clock Decoders) is at a low potential. In the selected decoder node 9 is pulled to Vdd by M9. Node 14 is pulled



DYNAMIC BIT LINE CLAMPS

FIGURE 7



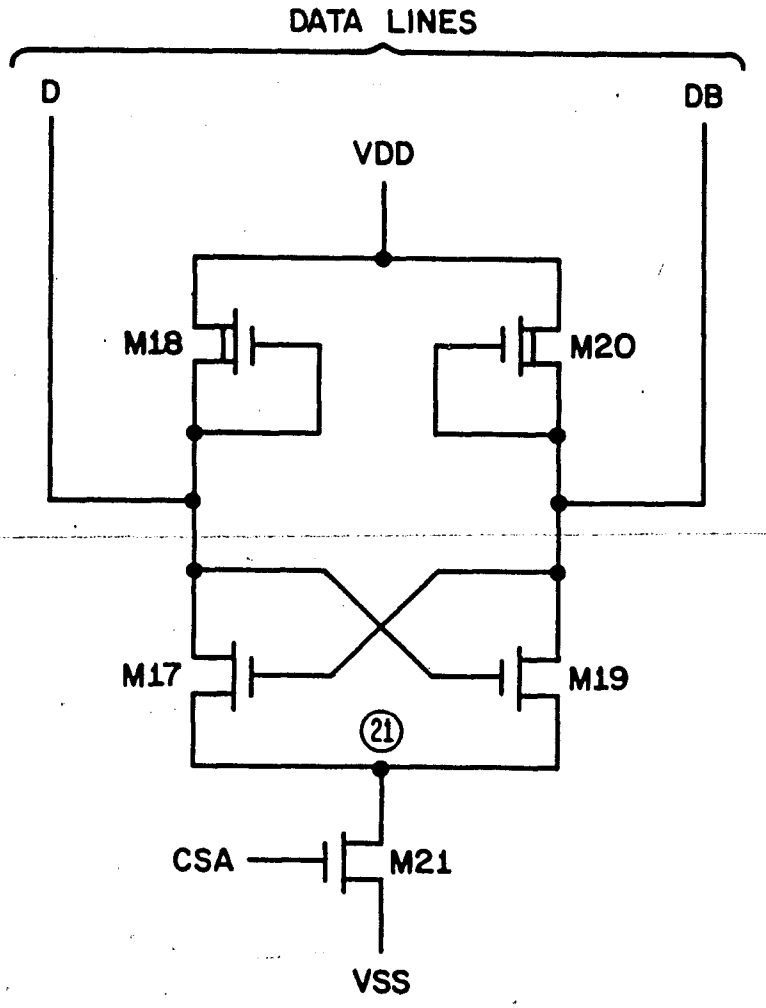
DYNAMIC DECODER

FIGURE 8

to one threshold below Vdd through M14, charging capacitor C15 which is the gate capacitance of M15. After sufficient time has elapsed for this precharge operation to occur, CD changes to a high potential. At this time M14 is at the point of turn off, so it cannot provide a discharge path for C15. The potential of node 14 rises with CD, turning M14 off hard. C15 partially discharges to charge C16, the other capacitance that exists on node 15, but enough charge remains on C15 to keep node 14 more than one threshold voltage above CD. Thus as CD changes to Vdd potential, M15 stays turned on and the word line is also driven to Vdd potential.

Only one device is necessary on the entire memory to pull up on CD. By delaying this pullup and using the precharged, bootstrapped decoder, we have replaced a multitude of "M13" devices by a single pullup device. The extra power dissipated by the other M13's is eliminated.

A dynamic sense amplifier is shown in Figure 9. It bears a strong resemblance to the memory cell, with the pullup resistors R1 and R2 replaced by depletion mode MOS transistor pullup devices M18 and M20. If CSA (Clock Sense Amplifier) is high and M21 is turned on, the circuit behaves exactly like a memory cell, with two stable states. The transistors in the sense amplifier have a sufficiently high gain so that, if the cross-coupled amplifier is not in a stable state but is slightly predisposed to one state or the other, it will quickly go to that state.



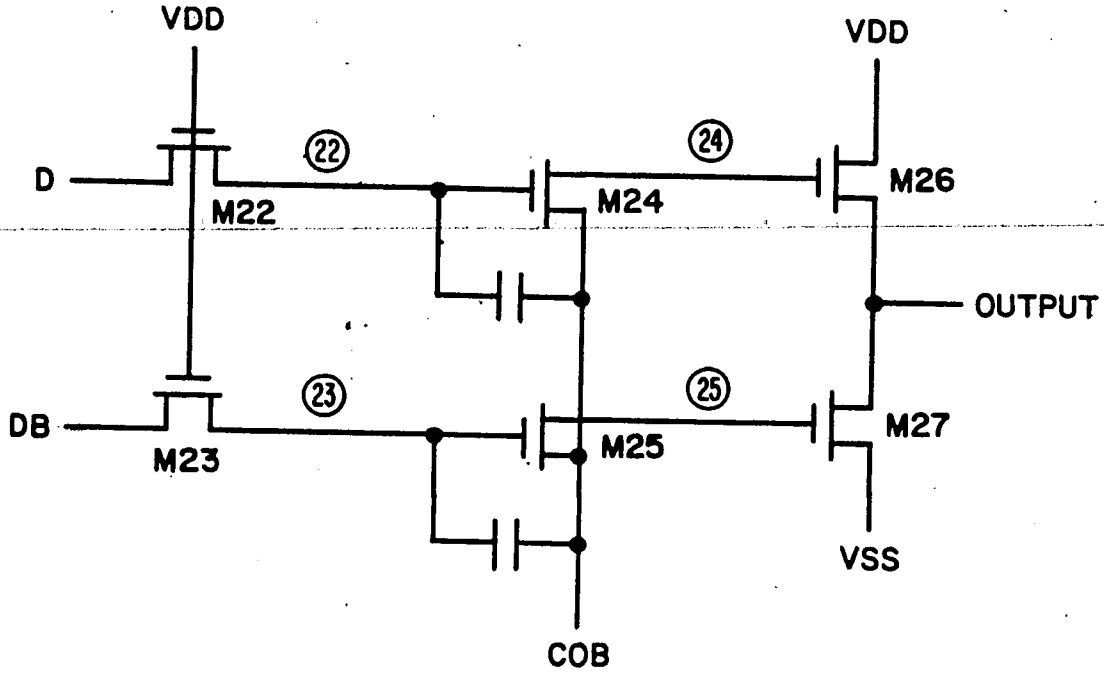
DYNAMIC SENSE AMPLIFIER

FIGURE 9

The sense amplifier increases the differential signal in the following manner. At the beginning of the access cycle, CSA goes to a low potential and the sense amplifier turns off. The data lines are pulled to a high potential by M18 and M20. After the decoders have been clocked, the selected word line is high and the selected row of cells begins pulling down on one of each pair of bit lines. One of the two data lines is pulled down by the bit line coupled to it by the selected column decoder. If a one is being read, the complement data line (DB) is pulled down. If a zero is being read, the true data line (D) is pulled down. Since it is the memory cell that is doing the pulling, this process is very slow.

However, after sufficient time has elapsed so that one data line is pulled slightly below the other (by perhaps two-tenths of one volt), CSA changes to a high potential. As the potential on node 21 decreases, the point is reached where the data line with the highest potential is one threshold voltage above node 21. Then the transistor (either M17 or M19) which pulls down the other data line begins to turn on. Thus the data line which was at a slightly lower potential is pulled down by node 21, resulting in an amplified differential signal.

A simplified output buffer (16) is shown in Figure 10. This output buffer performs in a manner similar to the decoders. Depending on whether a one or a zero was read, either node 22 or



DYNAMIC OUTPUT BUFFER

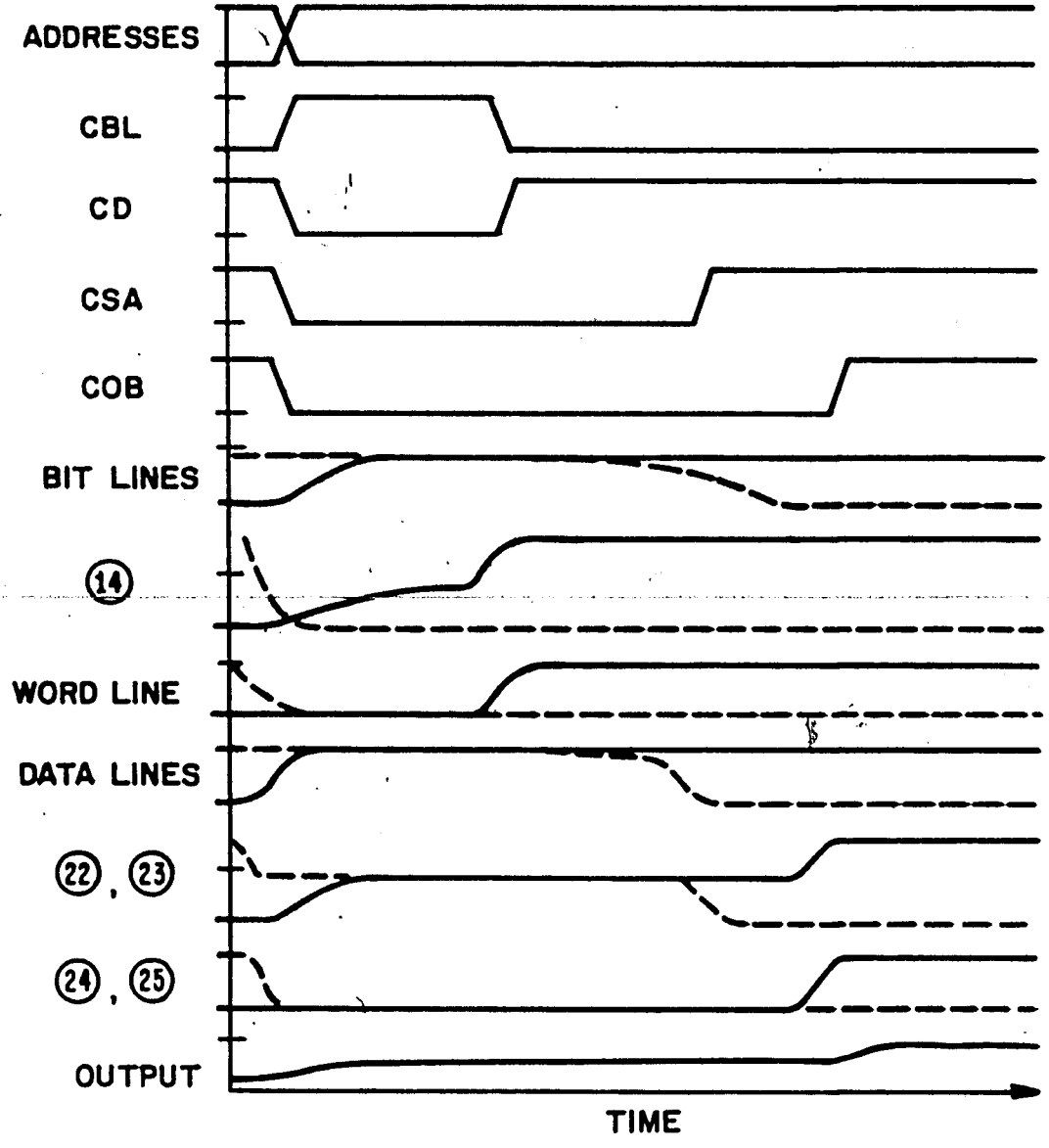
FIGURE 10

node 23 is precharged by D or DB. After this has occurred COB (Clock Output Buffers) changes from low to high, bootstrapping either M24 or M25 and driving either M26 or M27, outputting the data.

Figure 11 is a graph of the potential on various circuit nodes in the memory, as described in this section, versus time during an access cycle. In this graph, as in this section, the waveforms of CBL, CD, CSA, and COB are assumed. The next section deals with the general technique of generating these waveforms.

Clock Generation by Address Transition Activation. The following sequence of events must occur every time a new access cycle takes place. First, the new addresses are placed on the address inputs. Next, CD, CSA, and COB go to a low potential, and CBL goes to a high potential. Then, after the bit lines have recovered and the decoders have precharged, CBL goes to a low potential and CD changes to a high potential. Now some differential signal is allowed to develop on the data lines and then CSA goes high. Finally, after the output buffers have precharged, COB goes high.

The first event in this sequence is the development of the new addresses on the inputs. This implies that at least one address input has changed, and this change or transition is what triggers the generation of the clocks.



ADDRESS TRANSITION ACTIVATED MEMORY ACCESS CYCLE: READING A 1 AFTER READING A 0

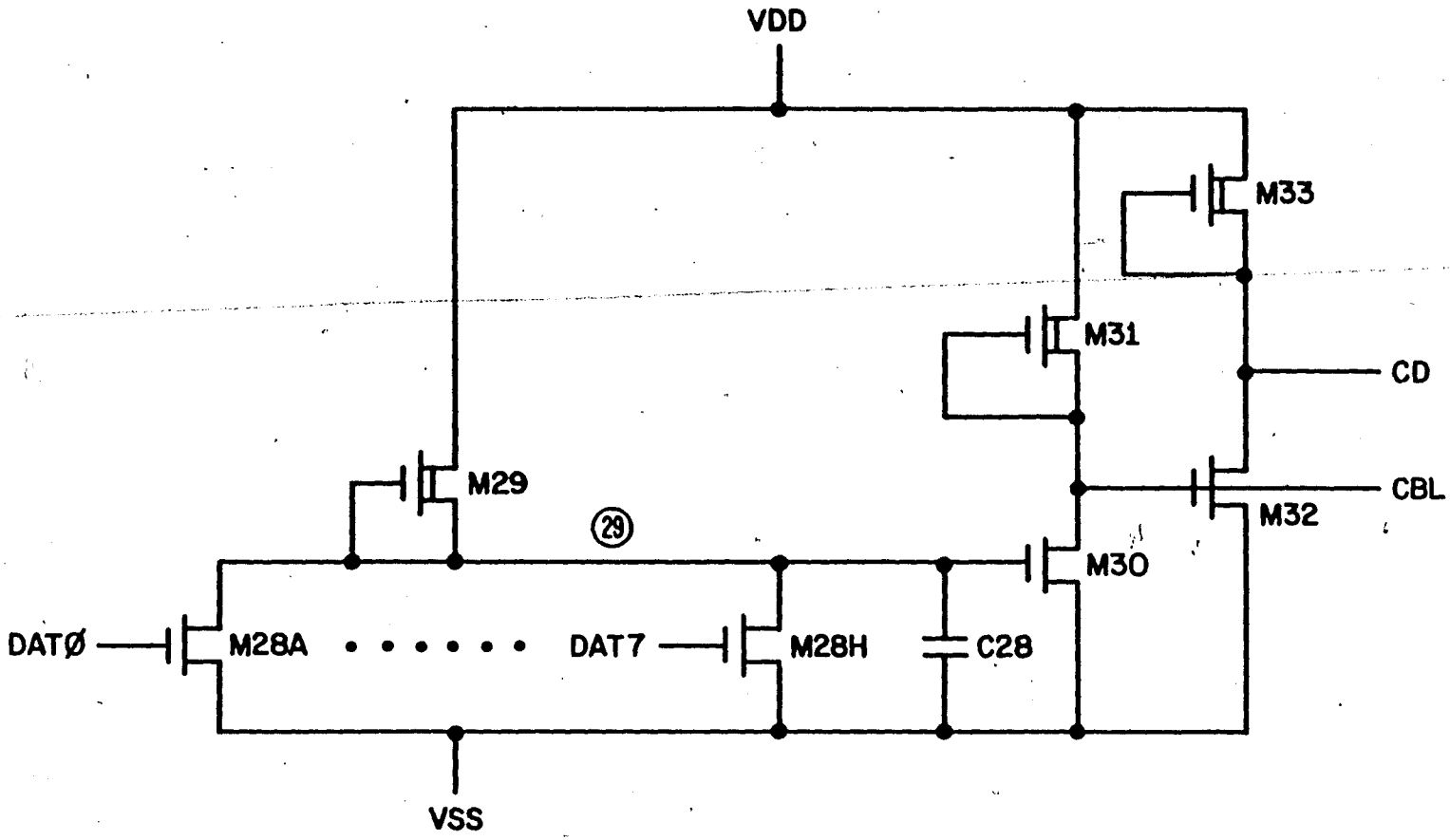
FIGURE 11

To each address buffer is connected a circuit called a transition detector. The design of this transition detector is investigated in Chapter Three. The function of the detector is to generate a narrow positive pulse, on a signal line called DAT (Detect Address Transition), whenever the address it is connected to changes from one state to another.

The DAT signal feeds into a chain of clock circuits which generate the clock signals. The first clock circuit, which generates CBL and CD, is shown in Figure 12. The operation of the circuit is as follows. When a pulse occurs on any of the DAT lines, corresponding to a change on any of the addresses, the M28 connected to that DAT line turns on momentarily and discharges C28. M30 is turned off and CBL goes high. M32 turns on and CD goes low.

M29 is a sufficiently narrow, long-channel device so that C28 is charged slowly. This charging rate is carefully arranged so that the bit lines are recovered and the decoders precharged just as the potential of node 29 gets to a threshold voltage. At this time CBL goes back low and CD goes high, just as desired.

Note that if another address changes before the precharge period is complete, C28 will be discharged again and the precharge period will be extended. The new decoder will have time to precharge, even though the bit lines are already recovered. The same thing applies to an aborted cycle - if the same address changes

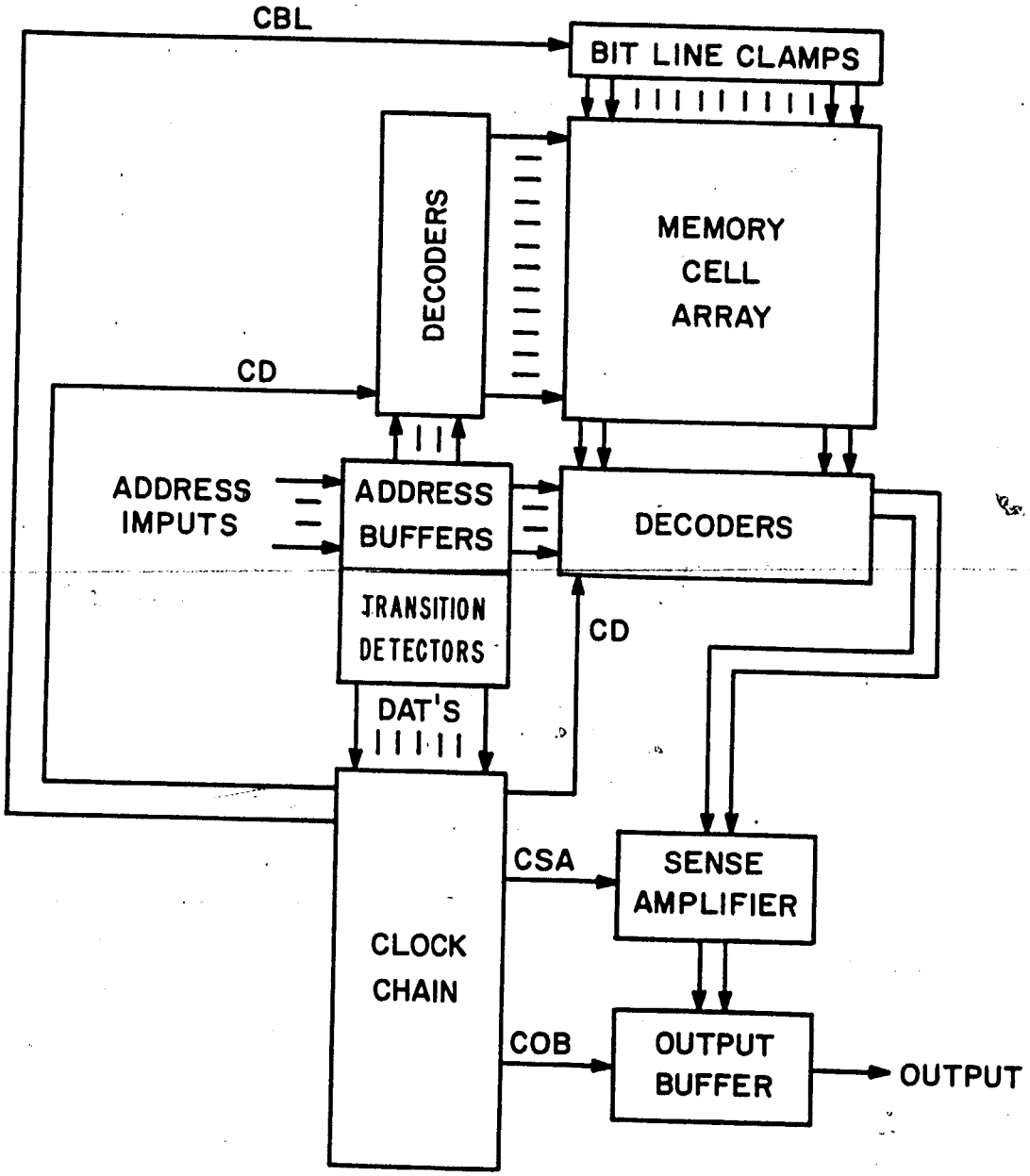


EXAMPLE OF A CLOCK

FIGURE 12

again, a new DAT pulse is issued and the precharge is extended. Making sure that the proper DAT pulse comes out under all of these conditions is the subject of the next chapter.

A complete block diagram of the Address Transition Activated Memory (except for write circuitry) is shown in Figure 13.



ADDRESS TRANSITION ACTIVATED STATIC MEMORY

FIGURE 13

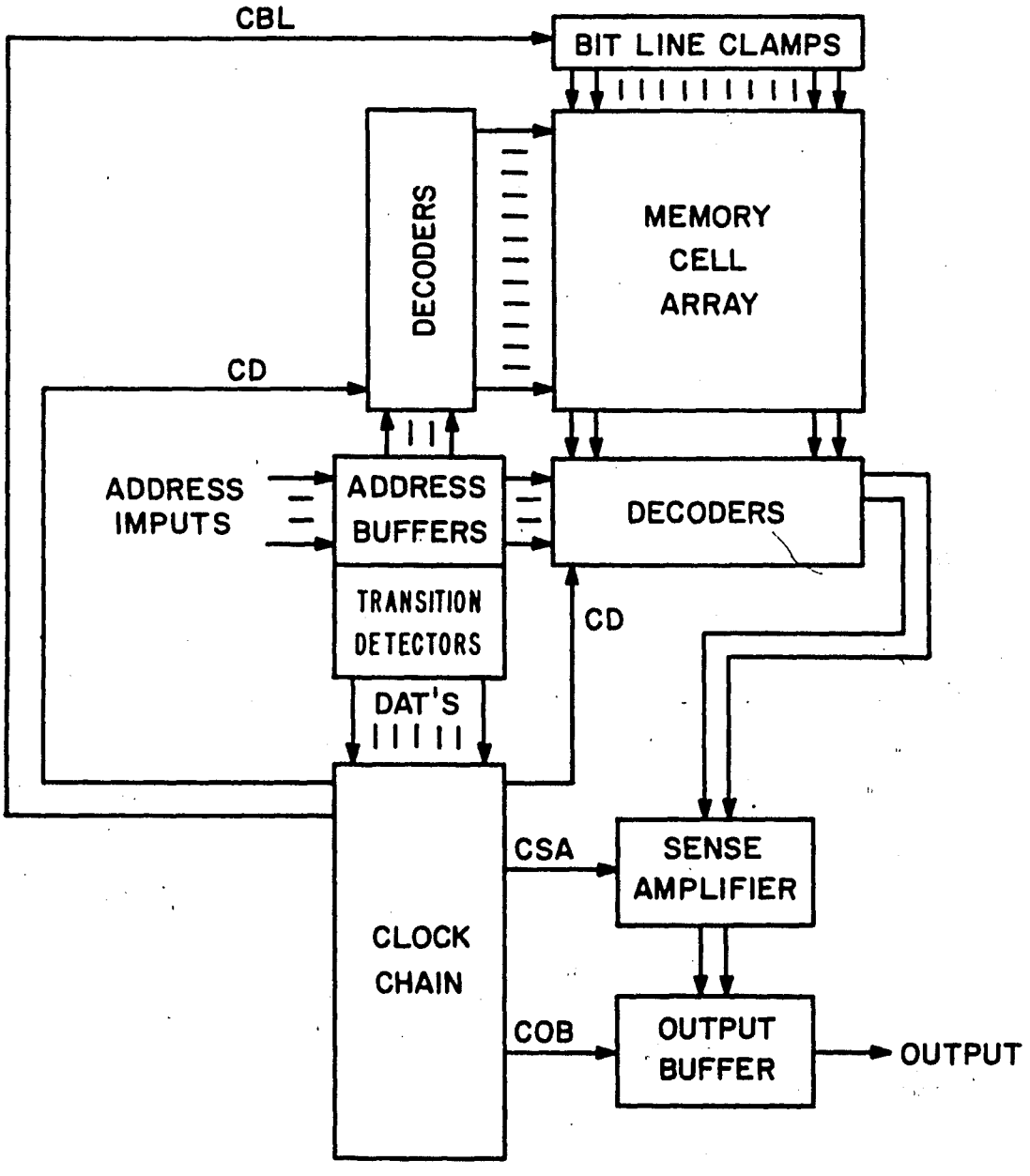
MICROMEX CORRECTION CODE (M-6)

CORRECTION

The preceding document has been re-photographed to assure legibility and its image appears immediately hereafter.

RP 30000

REMINGTON RAND
OFFICE SYSTEMS DIVISION



ADDRESS TRANSITION ACTIVATED STATIC MEMORY

FIGURE 13

3. Transition Detector Circuits - Theory and Design

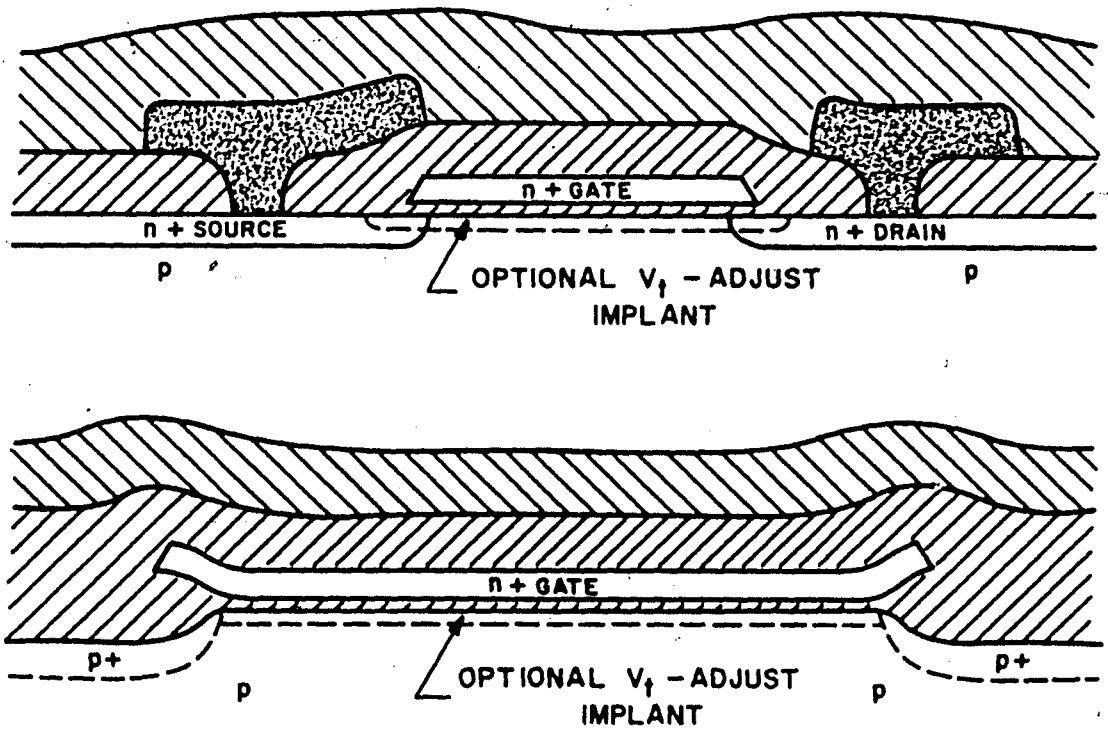
This chapter will detail the design of two transition detectors, an open-loop circuit and a closed-loop circuit. First, some additional constraints on these circuits, and some additional facts about their fabrication will be presented so that the ground rules for the designs are completely developed. A brief description of the procedure used to optimize and prove in a design will follow. Finally, the circuit designs will be presented.

The NMOS Technology. The circuits described in this work were designed to be fabricated with a silicon-gate n-channel MOS process (17) on lightly doped (5 ohm-cm) p-type wafers. This technology includes buried contacts between polysilicon and diffusion. It also includes optional threshold-adjusting ion-implantation steps to produce transistors with four different threshold voltages. Two of these threshold voltage (Vt) options have been used in circuits described in Chapter Two. These are the enhancement threshold and the depletion threshold.

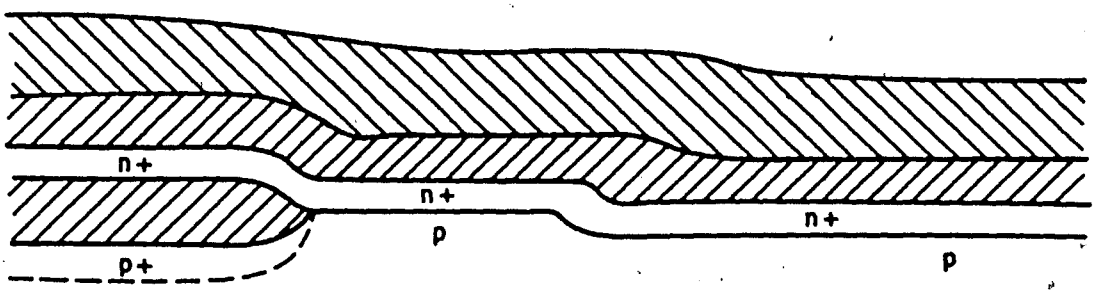
The other two ion-implant options produce a lower enhancement threshold and a light (more positive) depletion threshold. The application of these device types will be explained later in this chapter. Table 2 lists the nominal (desired) values for these threshold voltages as well as some other key parameters of the process. Figure 14 shows some of the key integrated structures produced by this process in cross-section.

Parameter	Nominal Value
L′ (Electrical Channel Length)	2.0um
tOX (Gate Oxide Thickness)	50nm
N (Substrate Doping Level)	$1.0 \times 10^{15}\ cm^{-3}$
VtH (High Enhancement Threshold)	1.15V
VtL (Low Enhancement Threshold)	.723V
VtN (Light Depletion (Negative) Threshold)	-.447V
VtD (Deep Depletion Threshold)	-2.7V

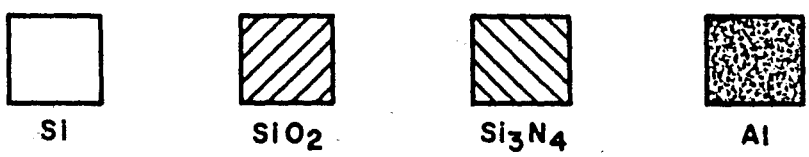
Table 2. Nominal Processing Parameters



NMOS TRANSISTOR - LENGTHWISE (TOP) AND WIDTHWISE SECTION (WITH SOURCE AND DRAIN CONTACTS)



BURRIED CONTACT BETWEEN GATE AND DIFFUSION

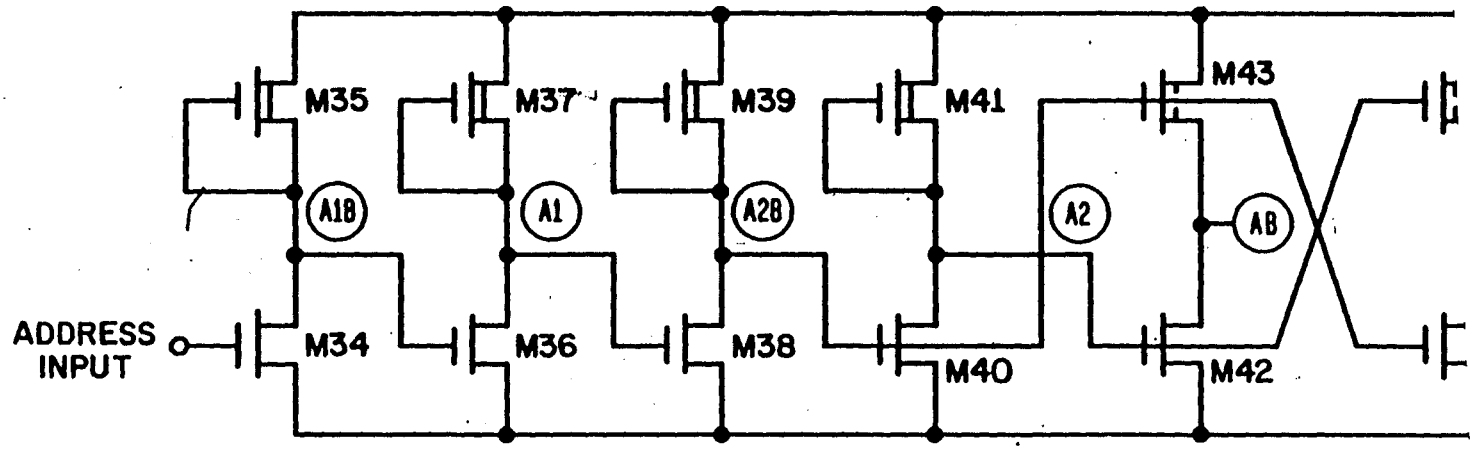


CROSS - SECTION OF VARIOUS STRUCTURES IN THE NMOS TECHNOLOGY

FIGURE 14

Layout Considerations. The overall layout of the memory integrated circuit must be considered when designing individual circuits. This consideration involves not only constraints but also benefits, such as fortuitous placement of various signal nodes which the designer can "tap off" for use in a particular circuit. At a minimum the transition detector must connect to the power supply leads, the address input and the DAT signal to the first clock. It turns out that other signals are available. These are node 29 in the first clock (Figure 12), and all of the circuit nodes in the address buffer. The address buffer is laid out as a unit with the transition detector by design, partially to ease the organization of the layout database and partially to allow sharing of circuit nodes. Node 29 is run to the transition detectors because this requires less space than running a separate DAT line from each transition detector to the first clock. The pulldown transistors on node 29 are actually incorporated into the transition detector layout.

A schematic diagram of the address buffer is shown in Figure 15. The node names A1, A1B, A2, A2B, A, and AB were chosen to represent the phase of each signal while giving it a unique name. The rest of the discussion disregards that there is more than one address buffer. The numbers 1 and 2 do not refer to different address leads.



ADDRESS BUFFER

FIGURE 15

The first four stages of the address buffer (M34 through M41) are simple NMOS inverters, in which the output is pulled high by the depletion pullup device, unless the enhancement pulldown device is turned on by a high on the input. These stages serve simply to amplify the TTL address input. The comlementary outputs (A and AB) of the address buffer are driven by push-pull stages in which the pullup is driven by a signal generated earlier in the circuit. This push-pull method saves some power (for a given delay) in driving the large capacitance associated with the address lines A and AB.

The pullup transistors in these final two stages are of the light depletion type. This is indicated by the broken bar between the source and the drain. These devices are operating in the source-follower (common drain) configuration; however, since the threshold voltage is negative, the output follows somewhat more positive than the input. This is of course provided that the pulldown transistor is turned off.

Additional Performance Constraints. The transition detector must respond to any possible transition on its input. A pulse on DAT must be generated whenever the address changes from one logic level to another. Allowance must be made for transitions that are very fast or arbitrarily slow.

Some margin must be allowed for the TTL input logic levels. This is so that a false detection will not occur due to noise on the

power supply busses, if the address input is held steady at a minimum level. For example, if a cell containing zero is being accessed, (or several cells containing zeroes in the case of a memory with a multi-bit word), a large spike of current in the VSS bus will occur when the output pulldown transistor turns on. This current is the discharging of the fifty picofarad load capacitance. This current must flow through resistance and inductance in the VSS lead, changing the potential of VSS on the chip. The transition detector will see this as a decrease in the level of the address input; if the input is a 1, the level will drop below the minimum input 1 level, perhaps by several tenths of a volt. This noise should not cause a DAT pulse.

Finally, the DAT pulse will be used to generate the CD signal to strobe the bootstrapped decoders. Decoder strobing will occur a fixed time after the trailing edge of DAT. To allow optimization of this time, it is desirable that the edge of DAT bear a fixed relationship to the signal on the address lines. Otherwise, the first clock will have to compensate for the shortest possible delay from the address lines to DAT, but the system designer will have to allow for the longest possible delay.

Design Procedure. The design of the transition detector circuits began with a block diagram or conceptual description either independently conceived or from the literature. The first step in a design is to convert the block diagram into an NMOS circuit, using some (usually well-known) implementation for each of the blocks.

During this implementation various decisions must be made, such as whether to use static loads or source follower pullups. Usually such decisions were made in favor of a simpler circuit, unless significant power savings could be obtained. Simplicity was desired to reduce ciruit area and design time.

After the basic circuit topology is developed, sizing of the various devices must be decided upon. An initial guess at the proper sizes was taken based on experience and comparison of the circuit to other circuits implemented in NMOS. Also, an estimate of the parasitic capacitance associated with each circuit node was made, based on the sizes of the devices connected to that node.

Optimization of the circuit was carried out by performing simulations with the ADVICE (18) circuit simulation program. Repeated adjustments to device sizes and parasitic estimates were made and repeated simulations were performed until the design of the circuit appeared to be optimized. Some criteria for the circuit performance, in addition to those already mentioned, were: the width of the output pulses, the delay through the circuit and the sensitivity of the circuit to variations in manufacture.

These simulations took into account expected variations in the processing parameters, as well as allowed variations in operating conditions. This was accomplished by simulating the circuit at various combinations of processing and operational conditions which would minimize the performance of the circuit. These are called worst-case conditions.

The ADVICE program uses CFSIM, a curve-fitting short-channel IGFET model (19,20). This model represents the MOS device with a set of equations which were derived empirically from measured I-V characteristics of actual devices. The parameters of these empirical equations are related to processing parameters by another set of empirical equations. As a result of this empiricism, and also as a result of the desire for simple equations to achieve economy in the simulations, a certain amount of error is tolerated in the model. One particularly notable inaccuracy is the representation of a depletion transistor as an enhancement transistor with a negative threshold. The IGFET model neglects the sub-surface conduction path formed by the donor ions in the depletion devices (21).

Once a design was optimized, a layout was designed and machine coded using a computerized integrated circuit layout system. The computerized representation of the layout produced by this system was analyzed using a hierarchical circuit analysis program which extracted from the layout a topological description of the circuit and

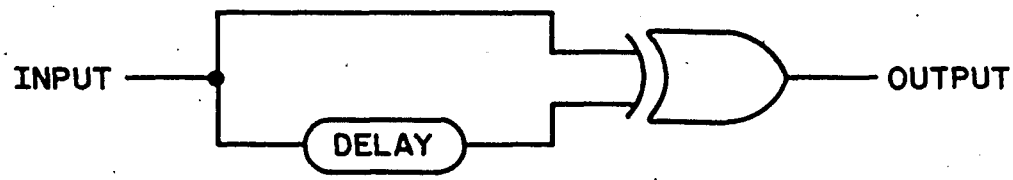
measured all of the device sizes. This information was used to verify the layout. This program also measured the area and overlap of all circuit nodes, allowing an accurate representation of the parasitic capacitance to be entered in the simulation.

Finally, a set of simulations using the accurate parasitic modeling was run in order to verify that the circuit met all performance requirements.

The Open-Loop Transition Detector. An open-loop transition detector is one which has no feedback paths. The open loop circuit has no logical memory; it resembles a combinatorial logic circuit. For D.C. conditions, its output is a fixed function of its input and does not depend on what state the input has been in previously.

A resistor-capacitor differentiator circuit suggests an open-loop transition detector (22,23), but a differentiator will not respond to an arbitrarily slow transition. The D.C. response of a differentiator is zero. A successful transition detector using a differentiator can be constructed only by including some feedback element, such as a trigger which uses regenerative feedback to generate a fast transition when its input crosses a threshold.

A commonly used edge detecting circuit (24) can be implemented in NMOS (25). This circuit is shown in Figure 16. The operation of the circuit is simple. If the input is stable, the output of the delay is equal to the input, and since both inputs to the exclusive



EDGE - DETECTING CIRCUIT

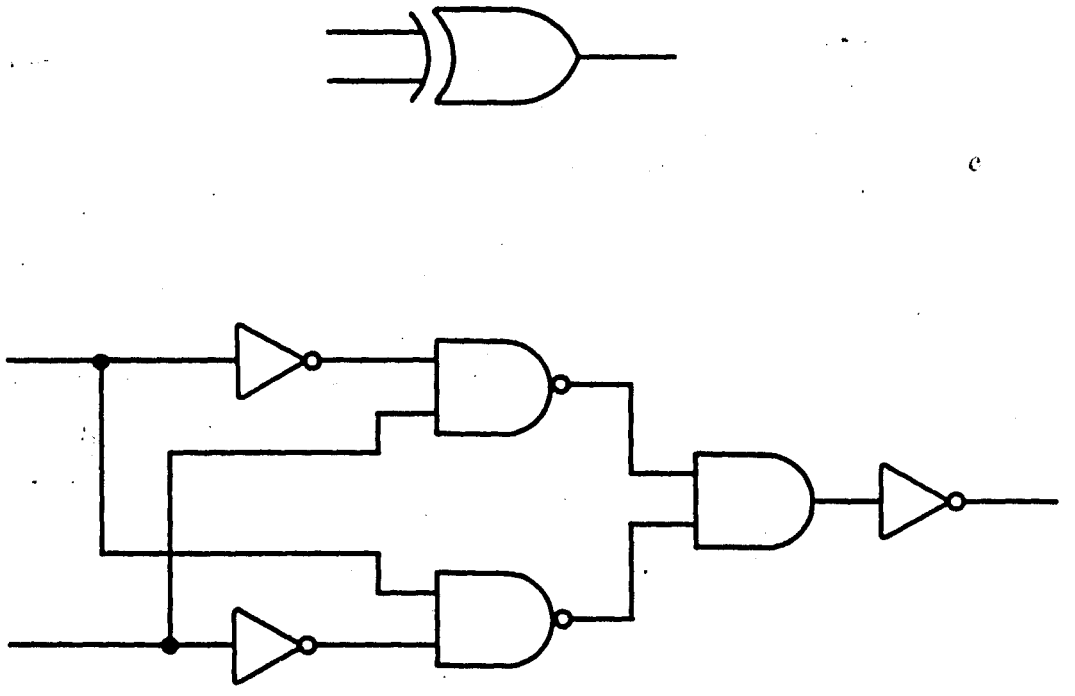
FIGURE 16

OR gate are the same, its output is zero. Now if the input changes from one state to another, this change will take some time to propogate through the delay. During this time the inputs to the gate are different. A pulse with width equal to the delay time is generated.

The decomposition of the exclusive OR gate shown in Figure 17 leads to an NMOS implementation. The address buffer is used for the delay, since it also provides the required inversions. The signals A1, A1B, A and AB from the address buffer can be directly connected to the NAND gates.

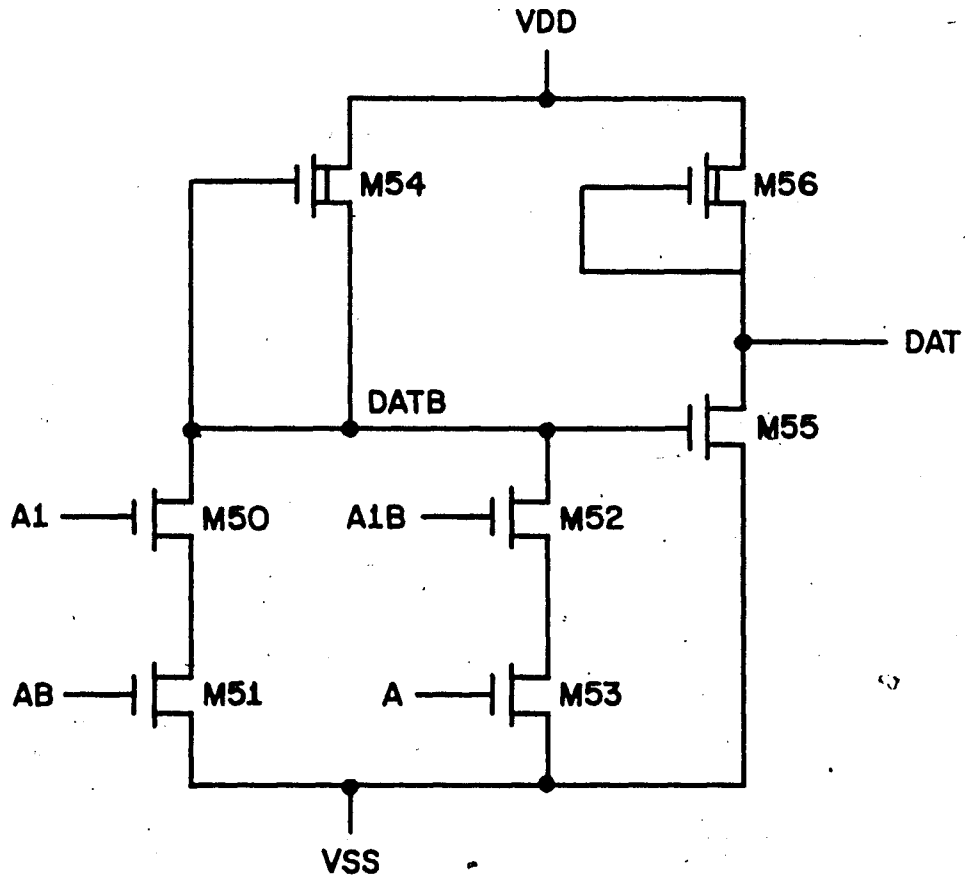
The resulting transition detector circuit is shown in Figure 18. The NAND gates are formed by the series connections of M50 and M51, and of M52 and M53. The outputs of these gates are dot-ANDed at node DATB. Final inversion of the output is performed by M55 and M56.

Let us examine the operation of this circuit in detail. First assume that the address input is initially in the logic zero (low) state. Then nodes A1, A2, and A in the address buffer will also be low, and A1B, A2B, and AB will be high. M51 and M52 in the transition detector will be turned on, but M50 and M53 will be off. Thus no current path will exist from node DATB to Vss, and DATB will be held high by M54. M55 will be turned on, holding DAT low.



DECOMPOSITION OF EXCLUSIVE OR GATE

FIGURE 17



OPEN-LOOP TRANSITION DETECTOR

FIGURE 18

Now let the address input make a transition to 1. M34 will turn on and A1B will be pulled low, turning off M36 and M52. M37 will pull A1 high, turning on M38 and M50. Now as M38 is pulling A2B low, M50 and M51 in series are pulling DATB low. M40, M43 and M44 as well as M55 will all turn off. M41 will pull A2 high and M56 will pull DAT high at roughly the same time. This is the beginning edge of the required DAT pulse.

The sequence of events following the transition is completed (in the address buffer and transition detector) as M43 and M44 turn on, pulling the address lines to their proper values. As AB is pulled below a threshold, M51 turns off, breaking the current path between DATB and Vss. DATB is recovered to Vdd potential by M54, turning on M55 and ending the DAT pulse. M54 is purposely made a long channel, narrow, low gain transistor so that the recovery of DATB to Vdd will be somewhat slow. This causes the width of the DAT pulse to be somewhat greater, which allows certainty that node 29 in the first clock will be properly discharged to Vss by the DAT pulse.

A similiar sequence of events occurs for a 1-0 transition. Initially A1, A2, and A are high, and A1B, A2B, and AB are low. M50 and M53 are turned on, but M51 and M52 are off. DATB is high and DAT is low.

After the transition, A1B goes high, turning on M36 and M52. A1 and DATB are pulled low at about the same time. DAT and A2B rise

roughly simultaneously, following which A2 falls. M43 and M44 now pull the address lines to their proper values, and when A falls below a threshold, M53 turns off, DATB rises and the DAT pulse ends.

Because the address buffer and the transition detector are located close together on an integrated component, we can depend on normal variations in the manufacturing process to effect all of the devices in the circuit the same way. Thus we can count on the risetimes and falltimes in the transition detector (except for the risetime of DATB) to always be at least as fast as the risetimes and falltimes in the address buffer; the address buffer is guaranteed not to get faster than the transition detector and cause the DAT pulse to be lost. Also, the low gain of M54 and the slow risetime of DATB will assure that the DAT pulse is always wide enough.

An ADVICE simulation of the operation described above, with nominal processing parameters, Vdd=5.0V, and room temperature is shown in Figure 19. This simulation verifies the basic operation of the circuit as described above, and it remains to show that the design meets all of the requirements for a transition detector.

We begin by examining the response of the circuit to a fast, high voltage transition. The danger in "overdriving" the circuit in this manner is that the increased gate voltage may speed up the circuit, reducing the delay in some stages. This should not adversly affect this circuit because only the first stage can be speeded up

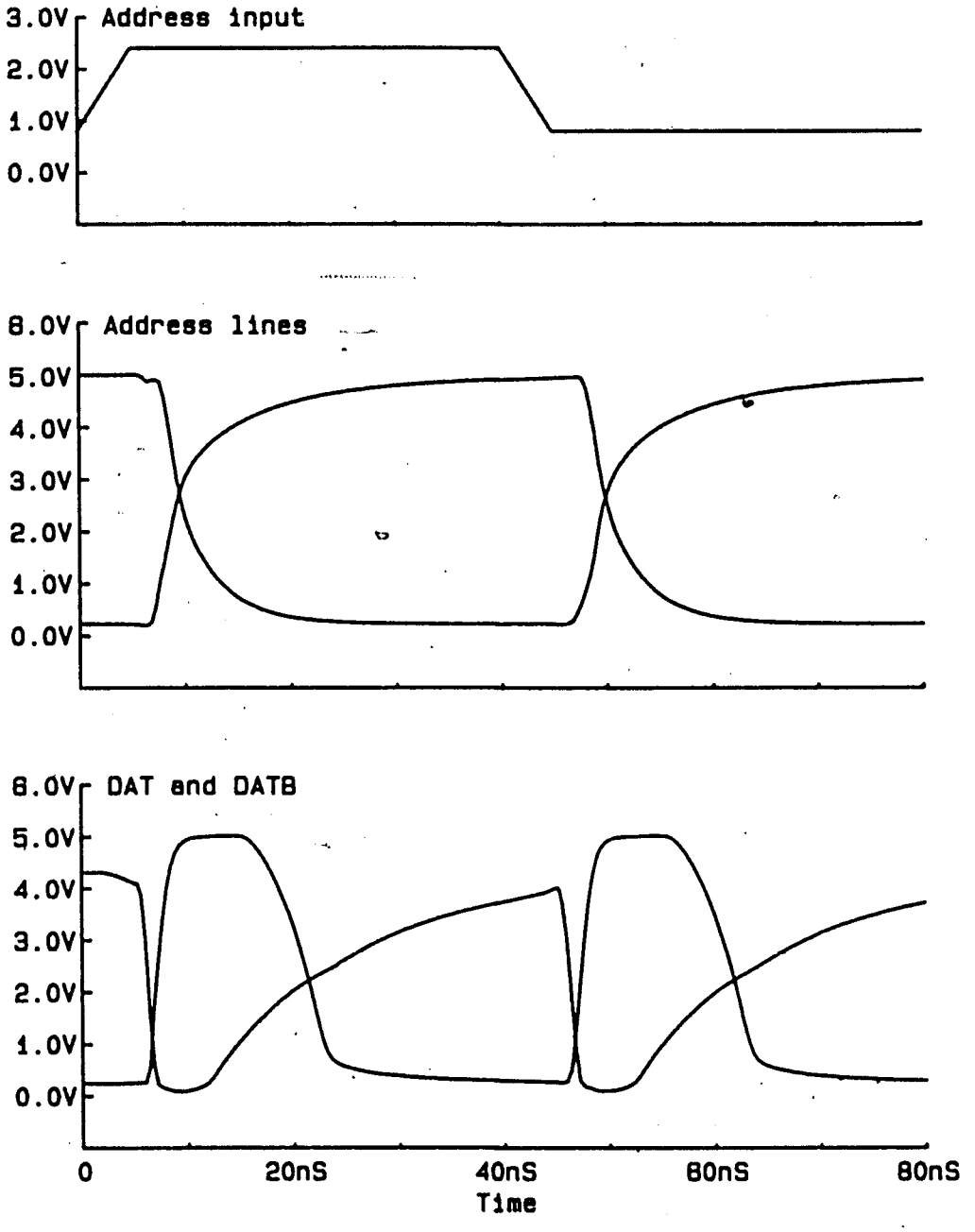


Figure 19. Simulation of open-loop transition detector with Vdd=5.0V and nominal process parameters

in this manner, if a 0-1 transition goes to a higher than usual voltage in a very short time. However, the operation of the transition detector does not depend on this stage.

Figure 20, Figure 21, Figure 22, and Figure 23 show the response of the detector to such an overdriven input for four different sets of conditions. These are combinations of extreme values for Vdd (4.5V to 5.5V) with extreme values for the processing parameters. The processing parameters are all varied simultaneously in a direction to either speed up the circuit to its maximum, or to slow it down as much as possible. These combinations of extreme parameters are called "slow" or "fast", depending on what they do to the circuit response.

Figure 24 and Figure 25 demonstrate that the circuit has a D.C. response to the address input. This assures that the component will work for very slow address transitions. The low gain of M54 is responsible for this D.C. response. Consider the situation as the address input is slowly changing from 1 to 0, and the potential on A1B is slowly increasing. As A1B gets above a threshold it will start trying to pull down on A1 and DATB. The potential on each of these nodes depends on the gain of the pullup device relative to the resistance of the pulldown. The low gain of M54 causes DATB to be pulled down faster that A1, so DAT goes high before A1 reaches a level that, through A2B and A2, causes A to go low.

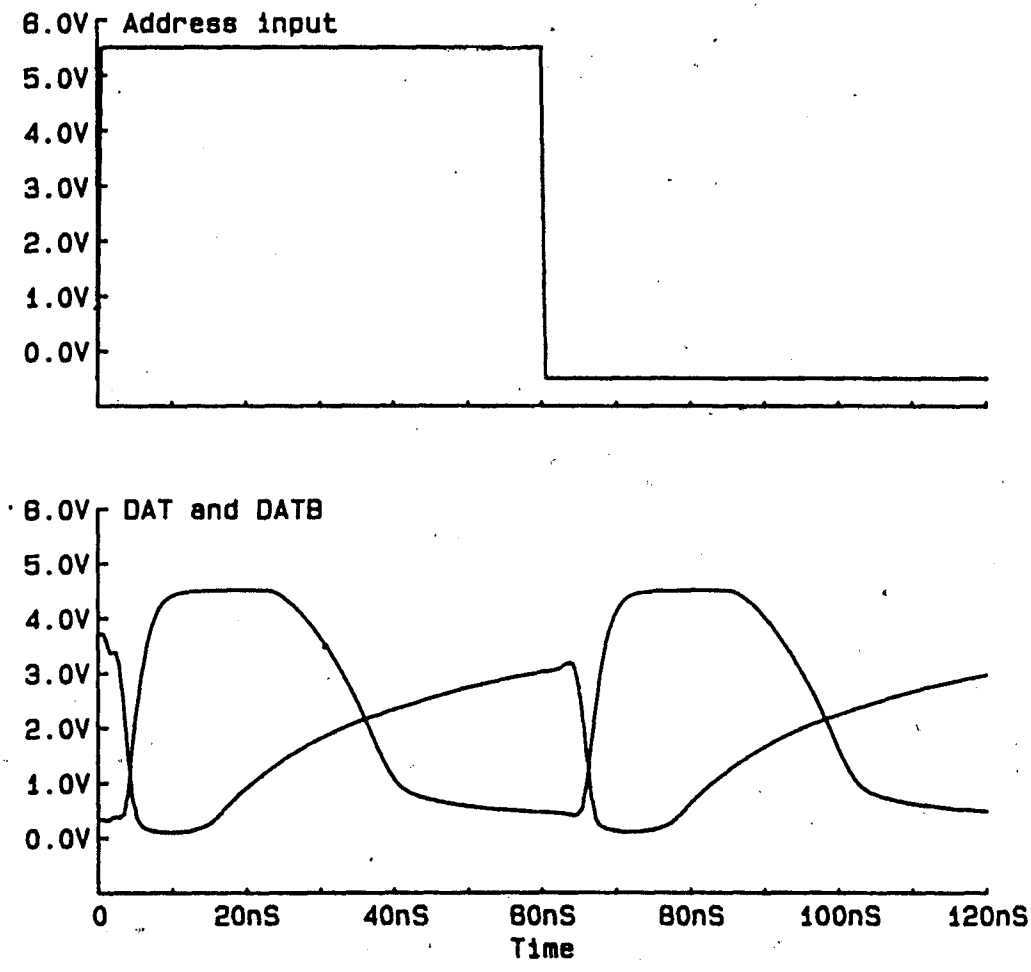


Figure 20. Simulation of open-loop transition detector with Vdd=4.5V and slow process parameters

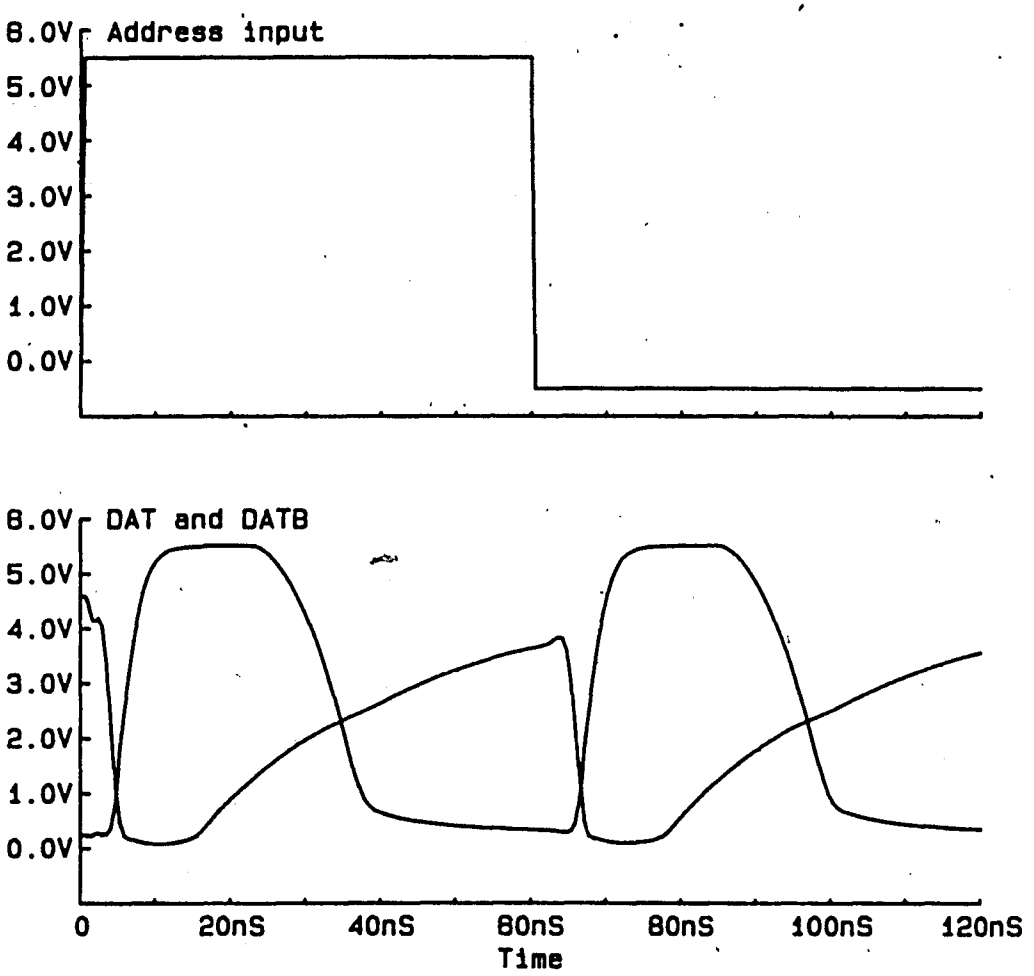


Figure 21. Simulation of open-loop transition detector with Vdd=5.5V and slow process parameters

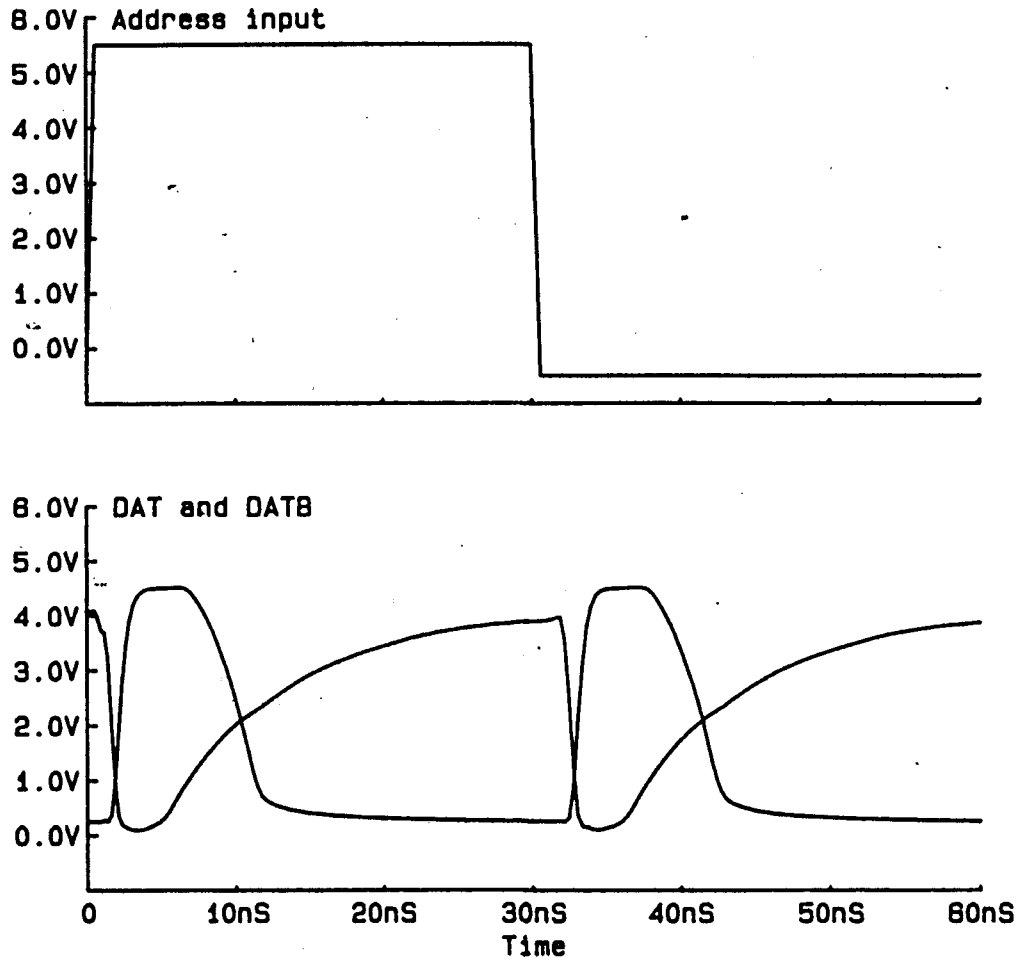


Figure 22. Simulation of open-loop transition detector with Vdd=4.5V and fast process parameters

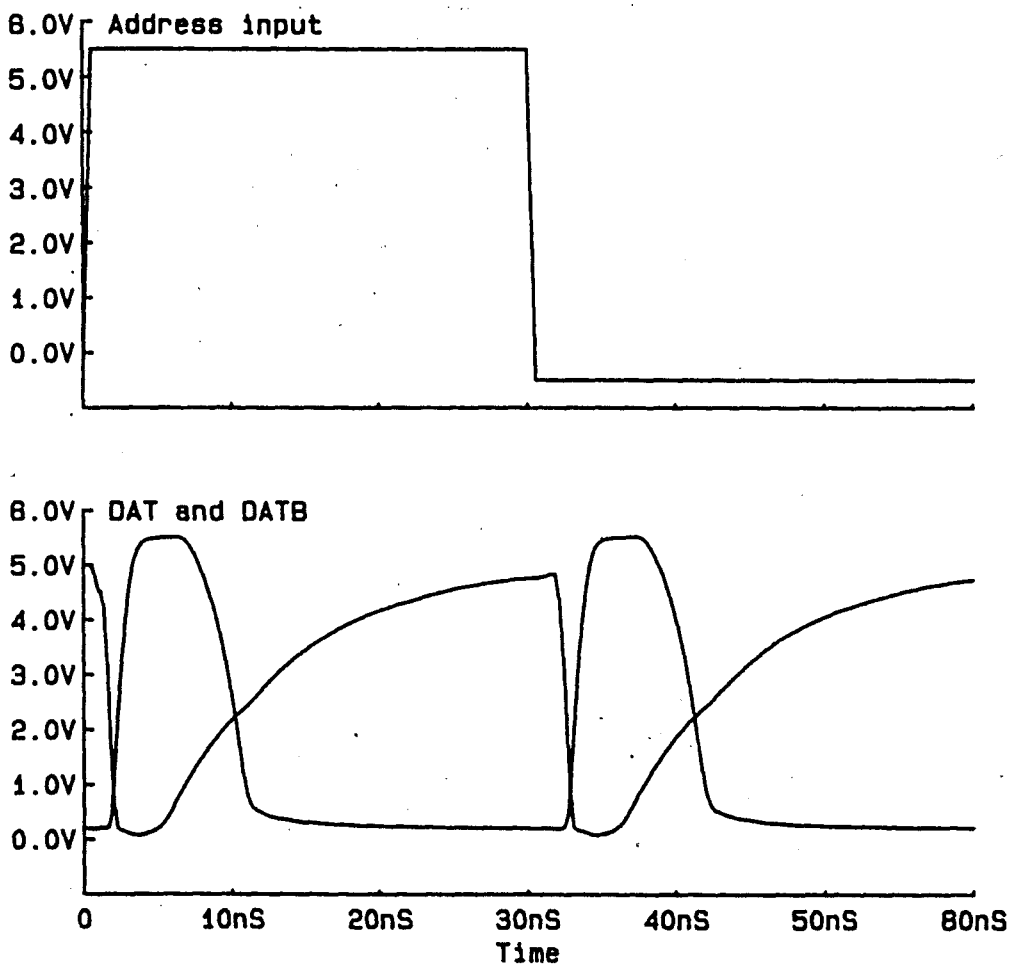


Figure 23. Simulation of open-loop transition detector with Vdd=5.5V and fast process parameters

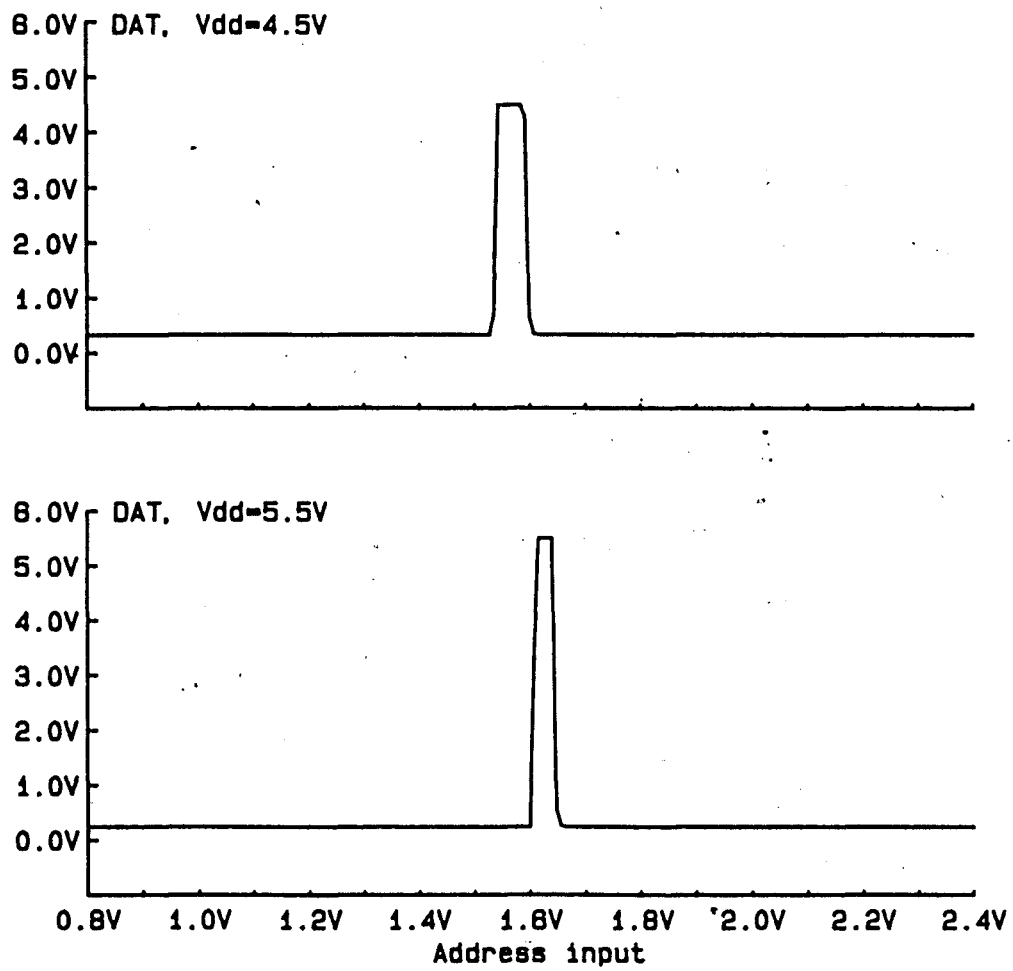


Figure 24. Simulated D.C. response of open-loop transition detector with slow processing parameters

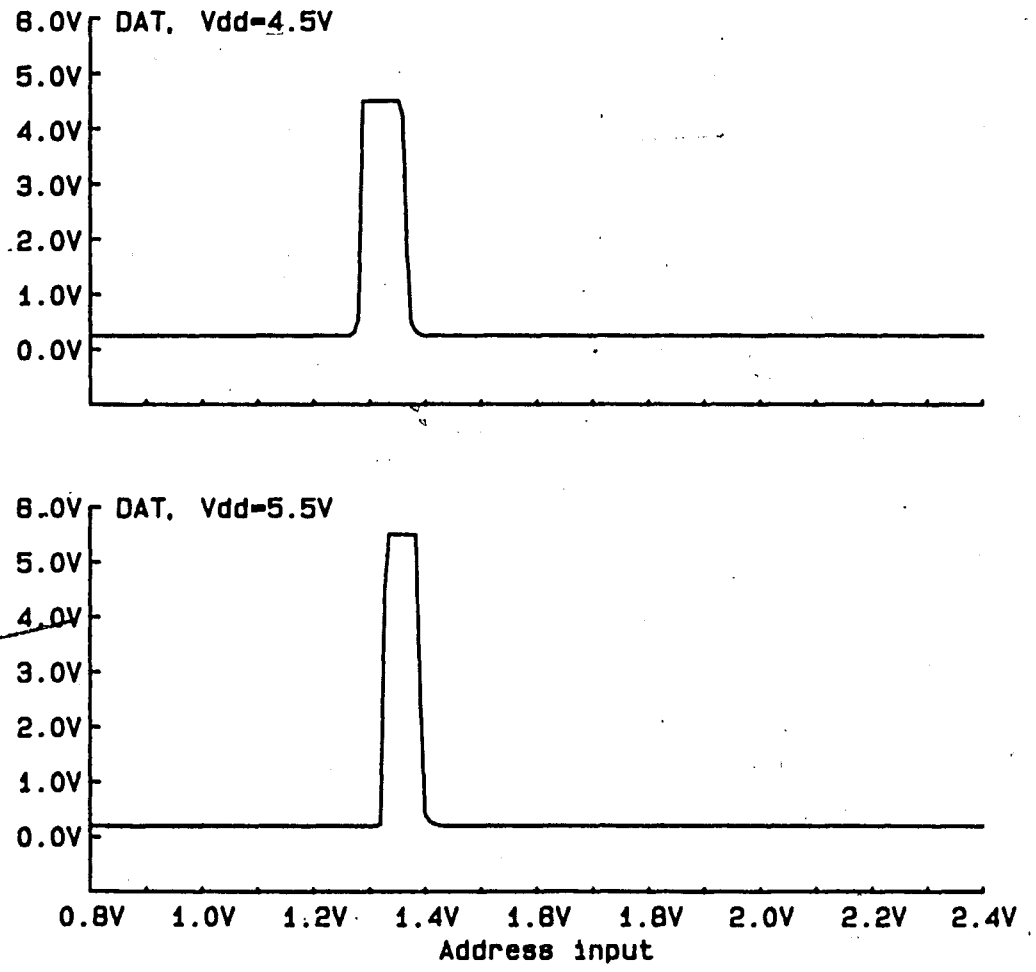
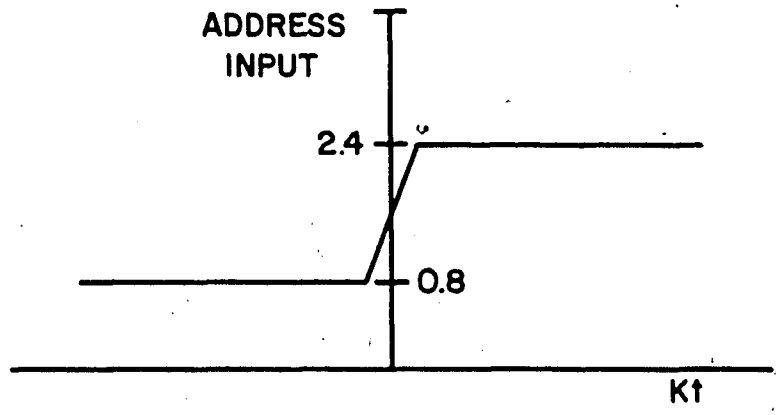


Figure 25. Simulated D.C. response of open-loop transition detector with fast processing parameters

In another region of the D.C. response of the circuit, A1 and AB are simultaneously high enough to pull down DATB. This is because A1 can pull down DATB at a potential at which it cannot pull down A2B, again because of the low gain of M54. That the D.C. response of this circuit includes conduction through M50 and M51 as well as conduction through M52 and M53 is a vital feature of this circuit. To understand this, consider the address input as a function of Kt, as shown in Figure 26, where K is an arbitrary constant and t is time. For positive K the function describes a 0-1 transition and for negative K the function describes a 1-0 transition. Now for K arbitrarily close to zero, the delays in the circuit caused by node capacitance have vanishingly small impact on the circuit response. The D.C. response is approached, and the bimodal (M50 and M51, M52 and M53) response occurs.

As K is increased in the positive direction, the propogation delays have an impact on the response. The delay from A1 to AB reinforces the M50-M51 response. Also, the delay from A1B to A causes the M52-M53 response to vanish, because M52 turns off before M53 turns on. Similiarly, as K becomes sufficiently negative, the M52-M53 response is reinforced and the M50-M51 response vanishes. If the D.C. response did not include both modes, a region of K might exist for which neither mode operated, and the transition detector would not be reliable.



GENERALIZED ADDRESS TRANSITION

FIGURE 26

In order to determine whether sufficient margins exist between the edges of the D.C. response of DAT and the TTL logic input levels, two special sets of processing parameters are constructed. These either maximize or minimize the capability of M34 to pull down A1B. The proper parameters are either slow or fast parameters, except that the depletion threshold is set to the opposite extreme than it would be set to for maximum or minimum speed.

A pair of simulations for these cases is shown in Figure 27. We see that sufficient margin exists, and that more margin exists for the minimum input 1 level than for the maximum zero level, as desired.

Finally, the issue of timing consistency must be addressed. Note that the trailing edge of the DAT pulse is brought about by an address line (A or AB) falling below a threshold. Due to the cross-coupled nature of the address line drivers (M43 through M45), we might expect that one address falls as the other rises with a generally uniform relationship. Thus we might also expect that the delay from the high-going address line reaching a given potential to the terminating edge of the DAT pulse might be fairly uniform; that is, it might depend only on the processing parameters and voltage levels that the rest of the memory will depend on and, in particular, the delay will not depend on the time between address transitions (in aborted cycle operation.)

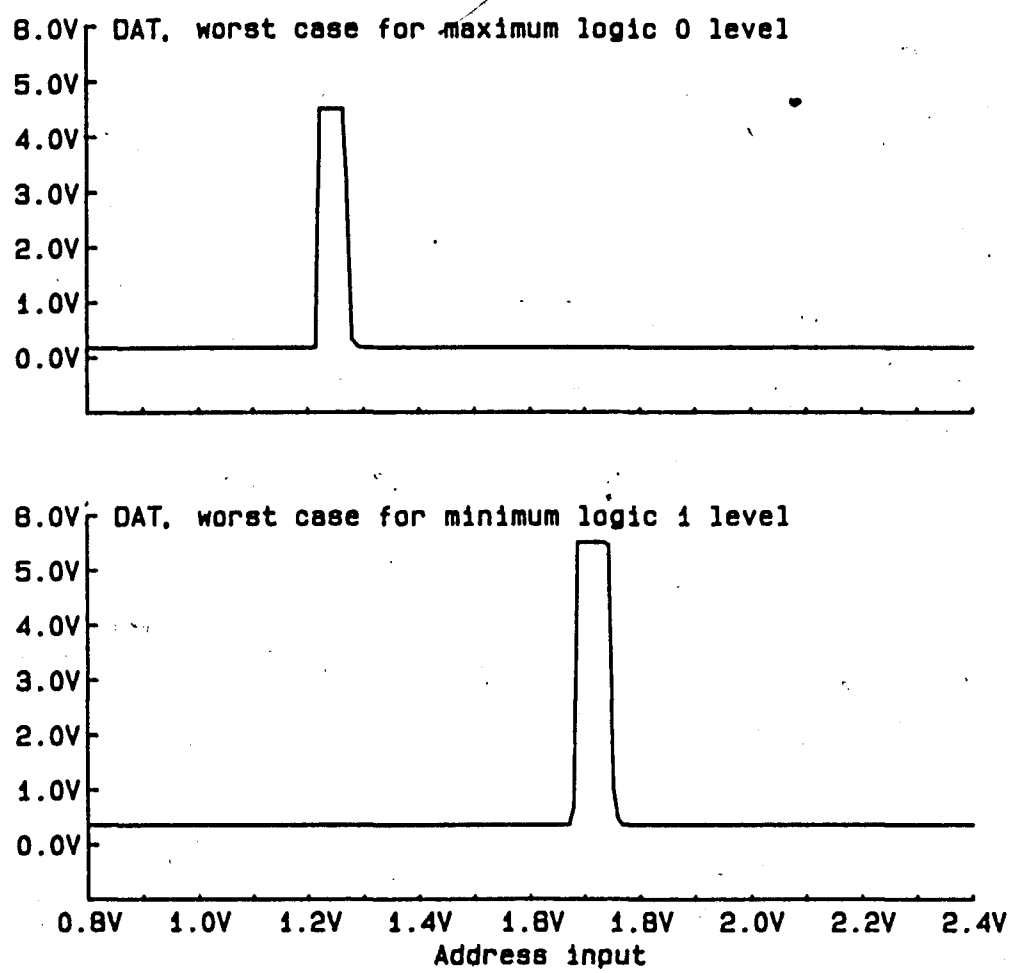
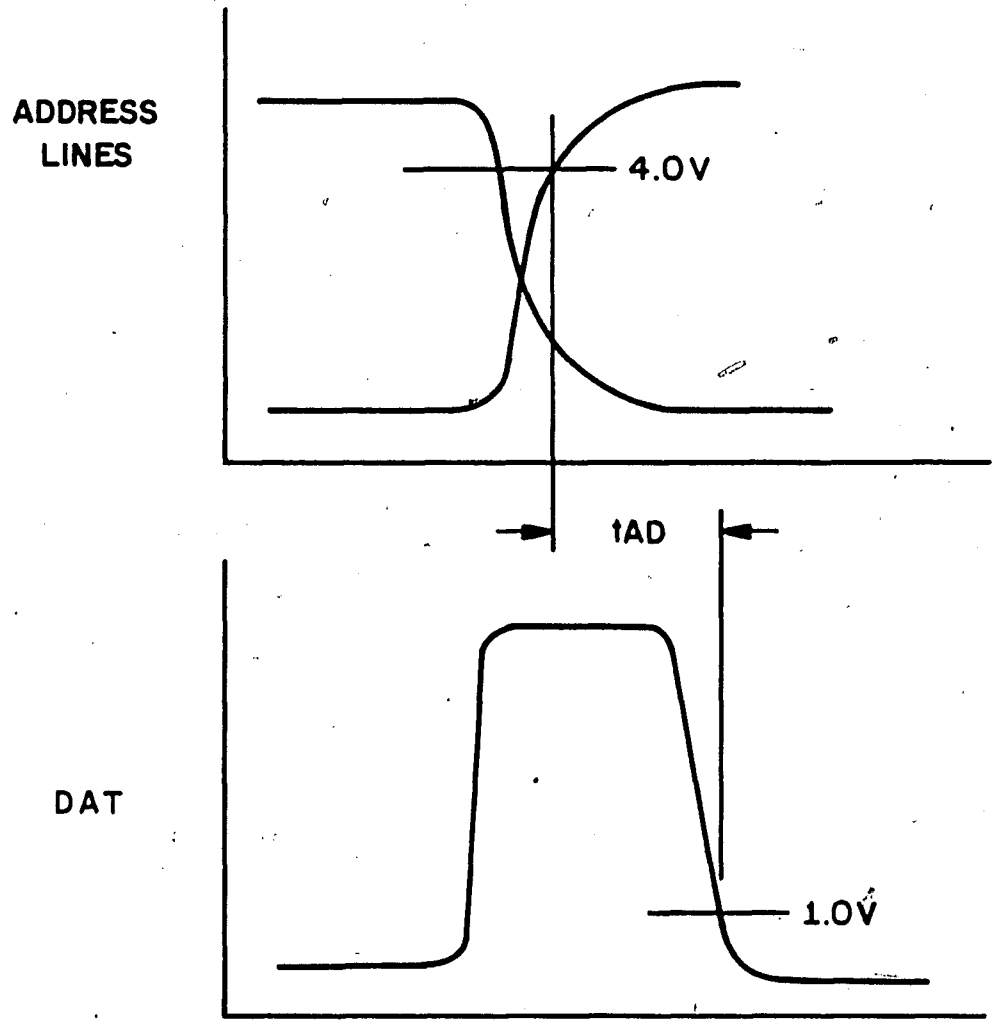


Figure 27. Simulated D.C. response of open-loop transition detector for worst case input level margins



DEFINITION OF tAD

FIGURE 28

This delay is named tAD and defined as shown in Figure 28 as the time from when the address line reaches 4.0V to the time when DAT falls below 1.0V. The time from one transition to the next, tTT, is defined as the time from the address input falls below 2.4V to the time it again rises above 2.4V. The Monte-Carlo capability of ADVICE (26) was used to perform several hundred simulations in which tTT was picked randomly and tAD was determined. Processing and voltage conditions were nominal for these simulations. Figure 29 shows a point for each of these simulations on a graph of tAD versus tTT. No points exist below tTT=8nS because DAT never rises above 1.0V for this narrow a "glitch" on the input. It is seen that tTT ranges from 7.4nS to 9.3nS, approaching 8.7nS asymptotically as the transitions become farther and farther apart. The 2nS variation in tAD is adequately small.

The Closed-Loop Transition Detector. One may argue that the open-loop transition detector might have a deficiency in that it might fail to detect a transition of some unforseen and unspecified form. It is not known how to prove that the open-loop detector must provide a DAT pulse for any possible transition, even though the preceding section has investigated a number of cases and made a very convincing argument. The closed-loop transition detector is an attempt to formulate a design which is provably reliable.

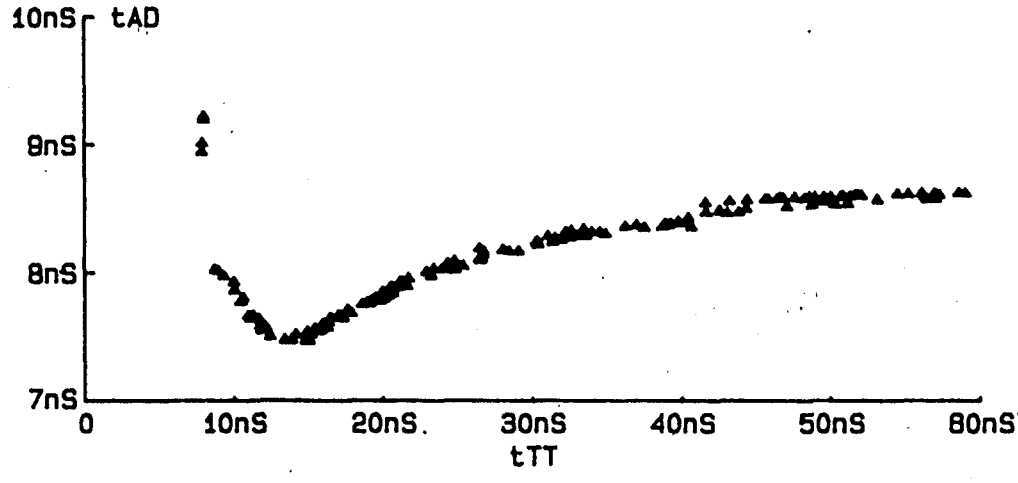
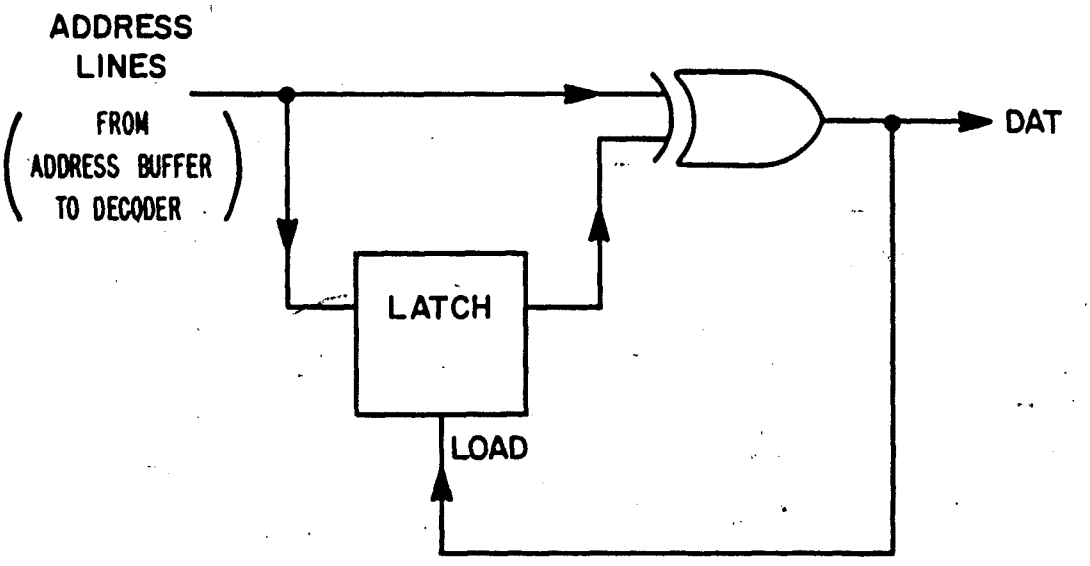


Figure 29. Simulation of open-loop transition detector showing delay from address line to DAT (tAD) versus time between address input transitions (tTT) for Vdd=5.0V and nominal process parameters

A block diagram of the closed-loop transition detector, as conceived by the author, is shown in Figure 30. The latch in this detector has the function of remembering the presently valid state of the addresses. If the address changes, the two inputs to the exclusive OR gate will disagree. When this occurs the DAT signal will become high. In addition to initiating the new cycle, the active DAT signal will load the latch with the new address information. When the latch contains the new address, the inputs to the exclusive OR gate agree again and the DAT pulse terminates.

Another configuration of the closed-loop transition detector has been accomplished in CMOS (27). This implementation has the latch inside the address buffer, so the address lines to the decoders come from the output of the latch. This scheme has the disadvantage of introducing additional delay into the address lines. However, it will turn out to have more compelling advantages.

The amount of time required to load the latch is a problem even in the configuration presented here, because it results in a much wider DAT pulse than is desired. This problem can be solved by eliminating some delay between DAT and the first clock. This is accomplished by replacing the exclusive OR gate with an exclusive NOR and connecting the output of it directly to node 29. The exclusive NOR gates of all the different transition detectors will now share one pullup transistor on node 29, which we will now call DATB. As



BLOCK DIAGRAM
OF CLOSED LOOP TRANSITION DETECTOR

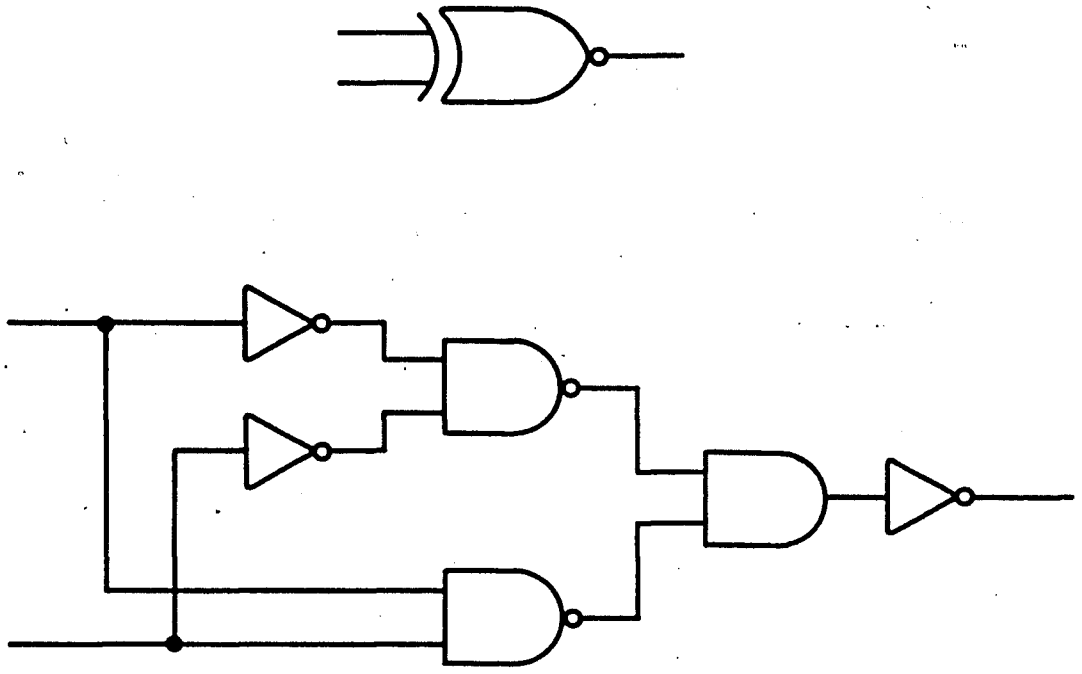
FIGURE 30

in the preceding section, however, we will discuss only one transition detector.

By reversing the true and complement of one of the inputs, the exclusive-OR gate of the open loop circuit becomes an exclisive-NOR. The decomposition is shown in Figure 31. The latch may be implemented as a memory cell, with the resistors replaced by depletion pullups and the LOAD signal acting as the word line. When an inverter is added between DATB and LOAD, the initial implementation of the closed-loop transition detector shown in Figure 32 is arrived at.

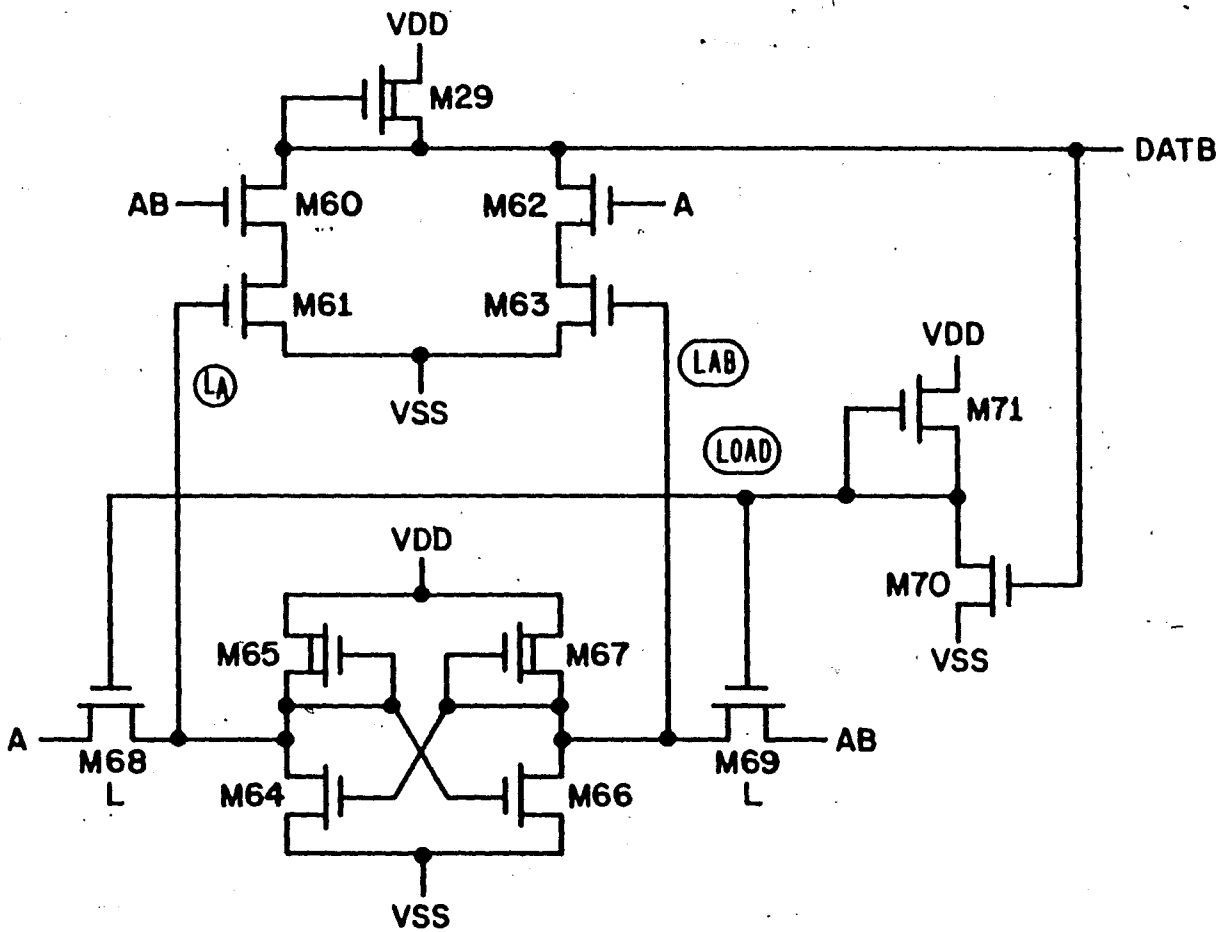
This implementation does not operate satisfactorily. During a transition when LOAD goes high, the node in the latch (LA, meaning Latched Address, or LAB) which was previously high is pulled low much faster than the low node is pulled high. As soon as the high node is pulled low, DATB goes high and LOAD goes low. The node which is now supposed to be high is left at an intermediate level, and it recovers to Vdd very slowly because the pullups M65 and M67 are very small (to conserve power.)

The result is that it takes much less time to load the latch at the next transition, and the timing of the detector is inconsistent. The problem could be remedied by increasing the gain of M65 and M67, but this would cause unnecessary power dissipation. Instead the modifications shown in Figure 33 were added.

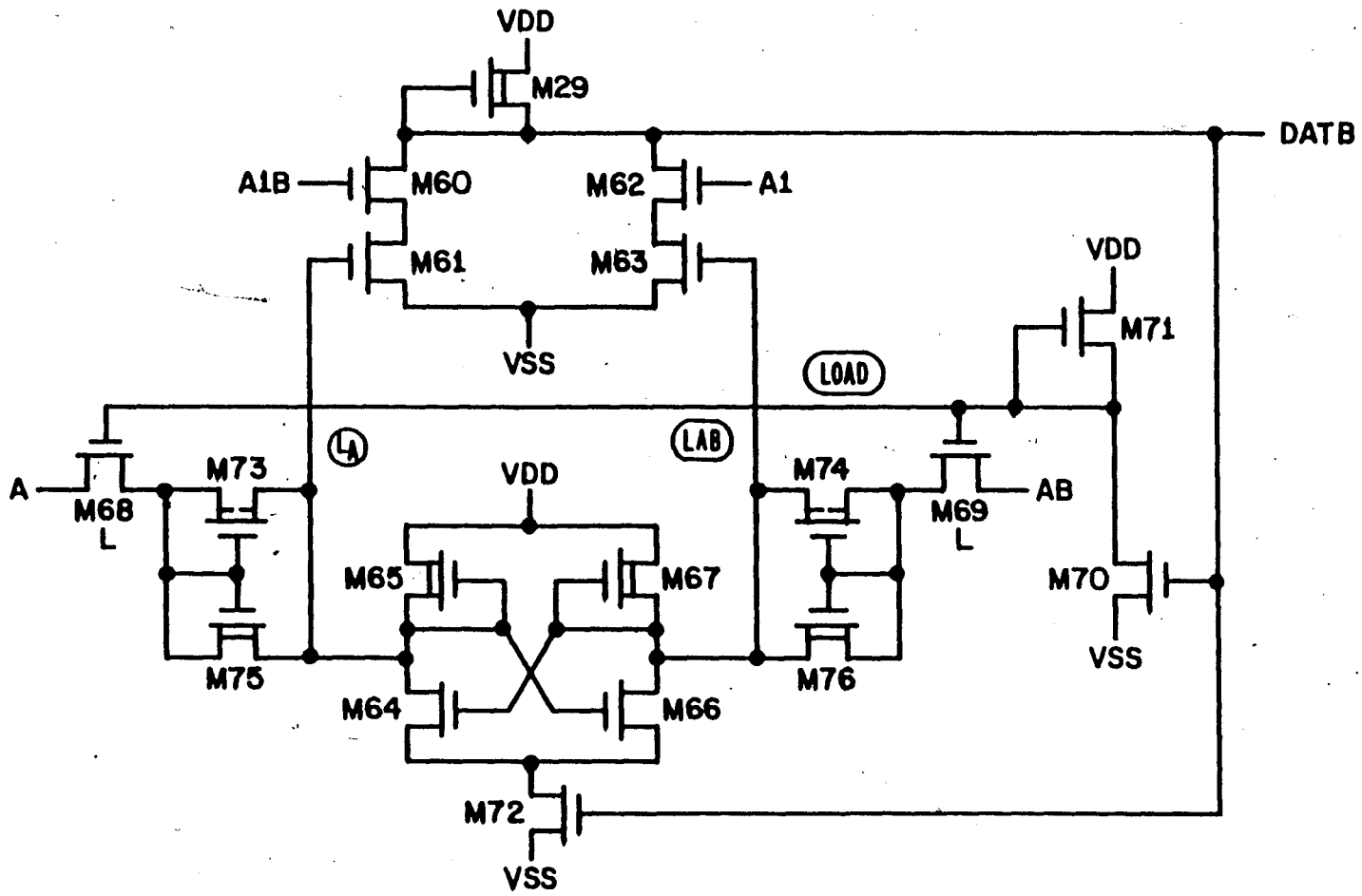


DECOMPOSITION OF EXCLUSIVE NOR GATE

FIGURE 31



INITIAL DESIGN OF CLOSED-LOOP TRANSITION DETECTOR
FIGURE 32



FINAL DESIGN OF CLOSED-LOOP TRANSITION DETECTOR

FIGURE 33

M72 disconnects the latch from Vss as soon as DATB goes low, allowing the low node to start pulling upward. M73 and M75 are added in series between M68 and LA to slow down the loading of the latch. When M68 is trying to pull LA high, M73 is turned on much harder than when M68 is trying to pull LA low. M75 is much smaller and simply provides a fine adjustment to the characteristics of the latch. Also, the upper inputs to the exclusive-NOR are connected to A1B and A1 instead of AB and A. This change allows LOAD to rise sooner, so that the latch is loaded faster when A and AB change. Optimization of this circuit was carried out by repeated simulation.

Let us examine in detail the operation of this circuit for a 0-1 transition. Initially, A ad A1 are low, and AB and A1B are high. LA must be low and LAB must be high for the circuit to be in equilibrium. DATB is high and LOAD is low.

A1B will go low and A1 will go high soon after the transition occurs. DATB is then pulled low, turning off M72 and M70. LA starts to go high but it does not get above a threshold voltage until A goes high. At this time AB goes low, pulling down on LAB. When LAB is below a threshold DATB is released and starts going high, pulled by M29. M29 is small to allow some delay for precharging the decoders, although some of this delay is now provided by the time required to load the latch. When DATB gets above a threshold, M72 turns on, applying power to the cross-coupled

inverters, and LOAD is pulled low, ending the loading of the latch. At this time LA is very close to Vdd and LAB is close to Vss.

This transition detector is guaranteed to give a DATB pulse whenever a transition occurs. Suppose that a transition has occurred but DATB has never gone low. Then LOAD must never have gone high, and the latch must contain the old address information. Then the inputs to the exclusive-NOR gate must be different, and DATB should be low. This inconsistency indicates that something is incorrect in the original supposition: either no transition has occurred, or a DATB pulse has occurred.

It is impossible to prove so easily that the open-loop transition detector is reliable. However, the elegant simplicity of the open-loop circuit has been traded for the ability to construct the proof. It will turn out that some performance has also been traded.

Figure 34 shows the inputs and outputs of the closed-loop transition for two transitions with nominal process parameters and Vdd=5.0V. Figure 35 shows some of the internal nodes, illustrating the operation explained above.

Next the response of the circuit to a fast, high-voltage transition is examined. Figure 36, Figure 37, Figure 38, and Figure 39 show simulations of this response for various combinations of Vdd and processing parameters. Two performance disadvantages of the closed-loop detector appear in these simulations. First, the

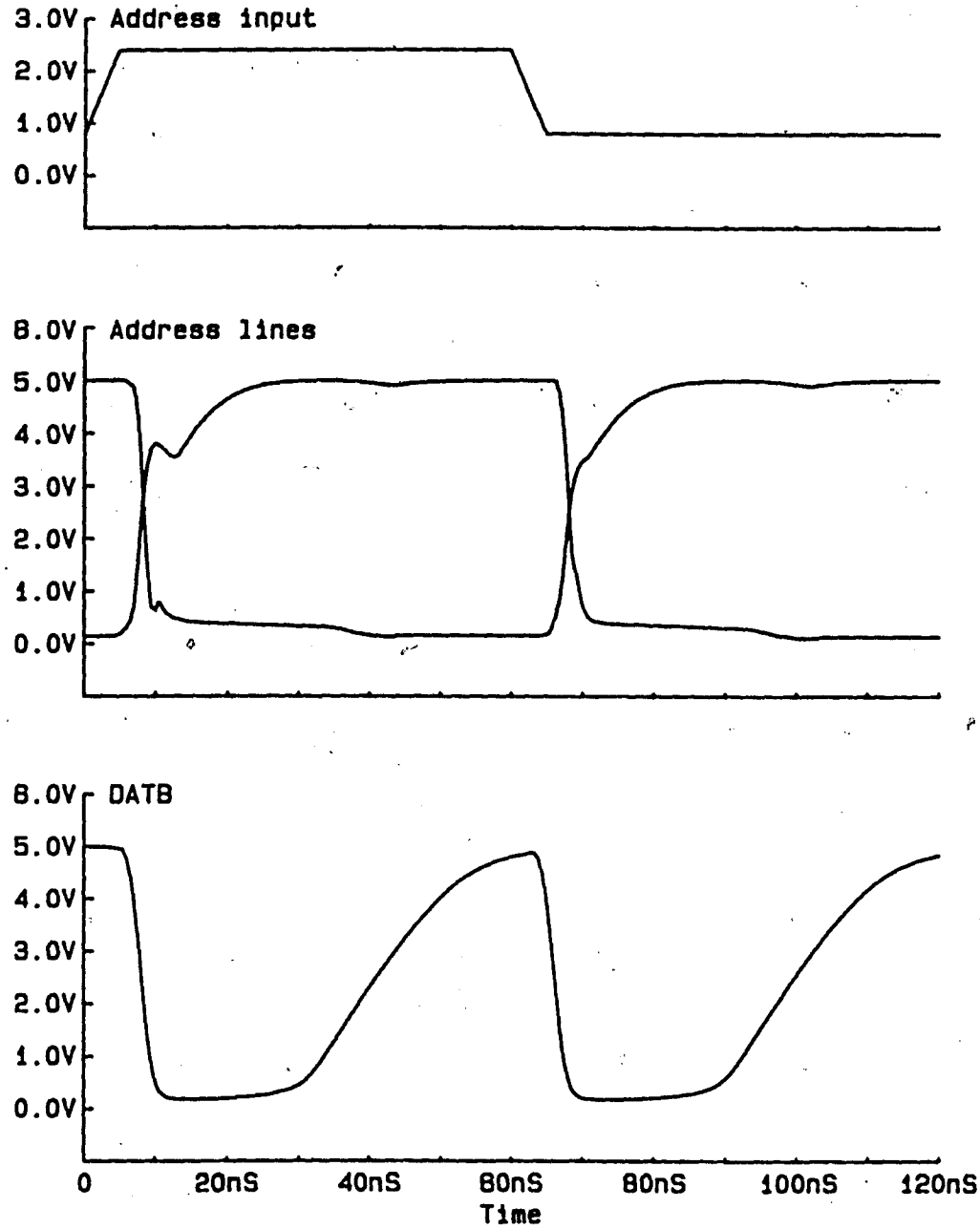


Figure 34. Simulation of closed-loop transition detector with Vdd=5.0V and nominal process parameters

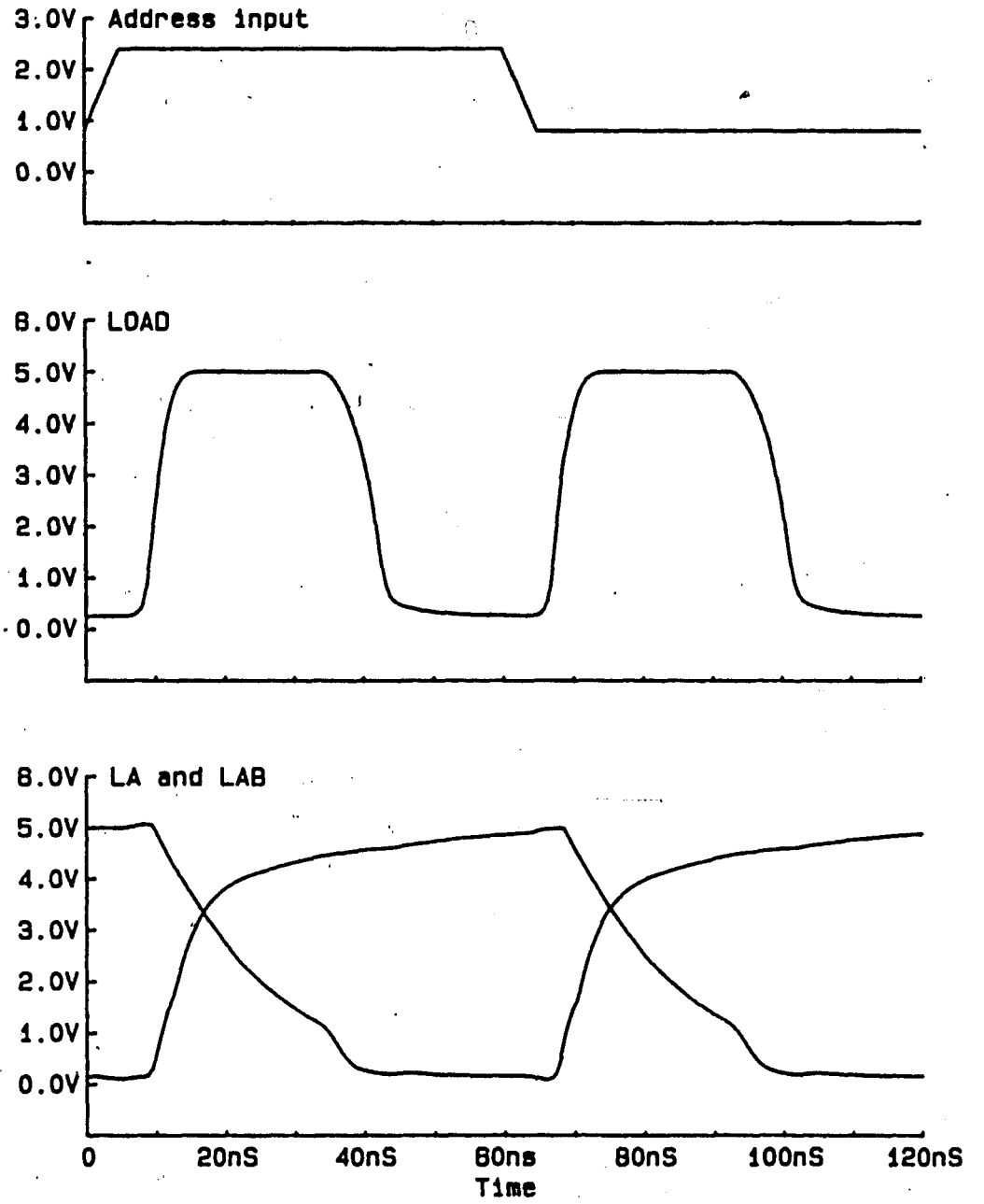


Figure 35. Additional nodes in nominal simulation of closed-loop transition detector

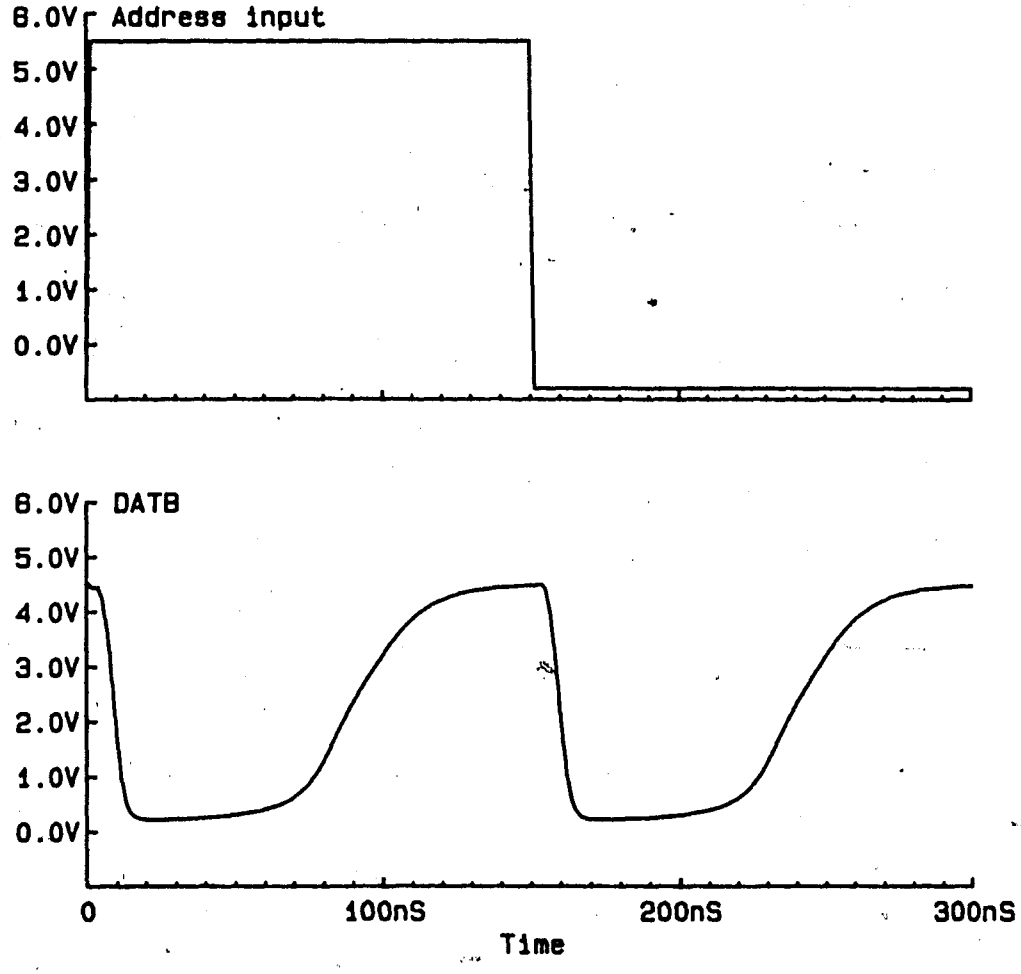


Figure 36. Simulation of closed-loop transition detector with Vdd=4.5V and slow process parameters

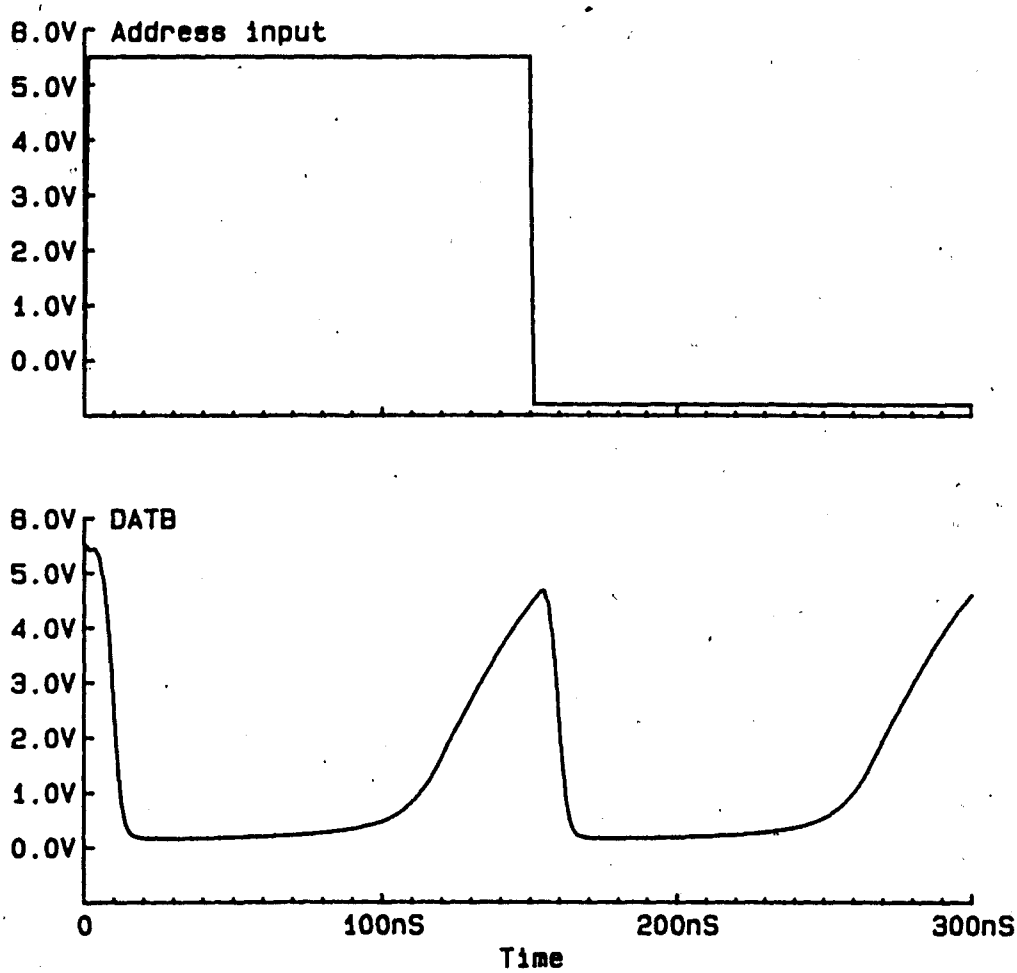


Figure 37. Simulation of closed-loop transition detector with Vdd=5.5V and slow processing parameters

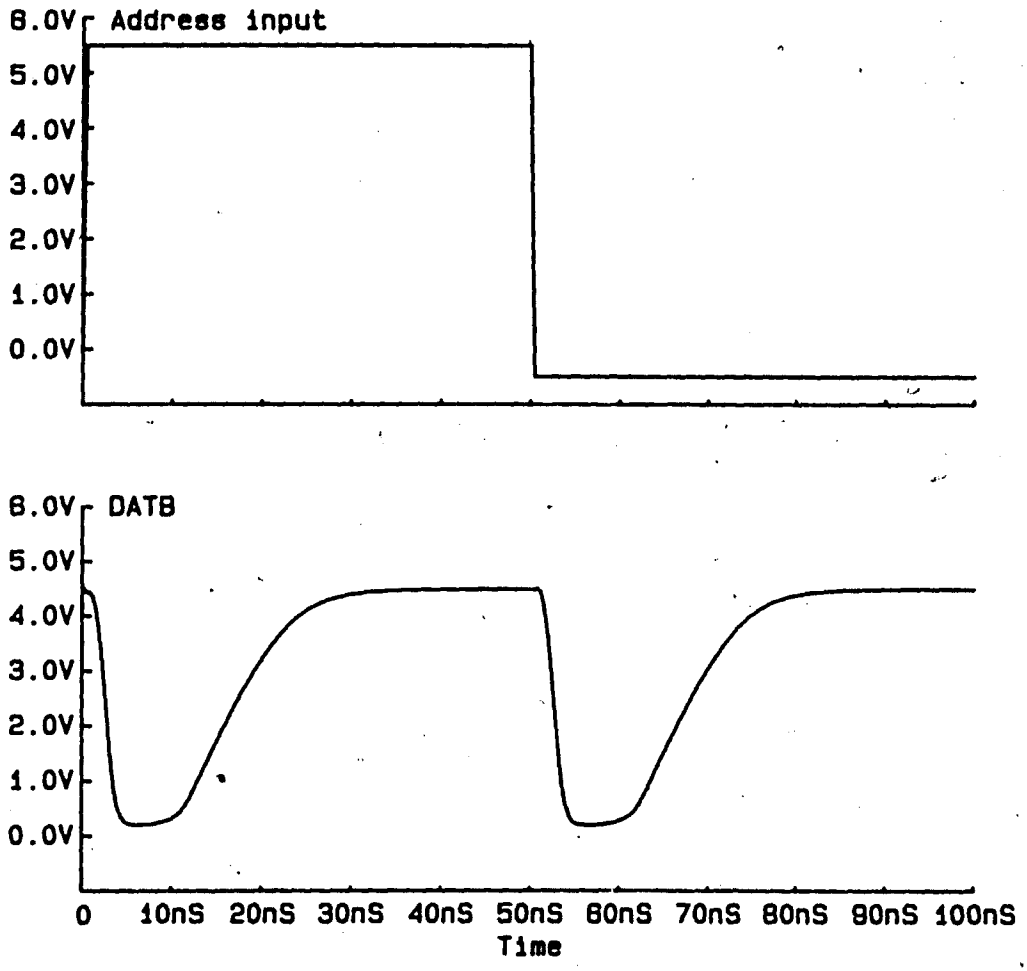


Figure 38. Simulation of closed-loop transition detector with Vdd=4.5V and fast processing parameters

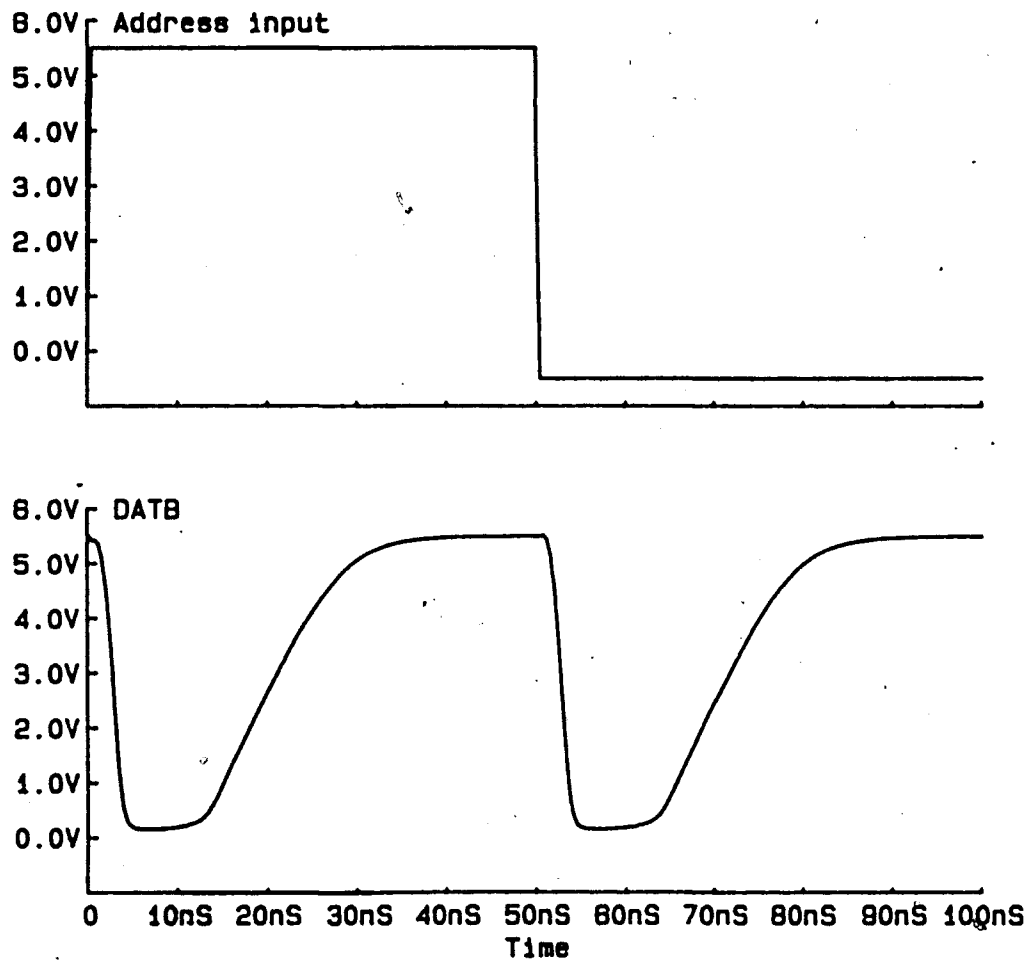


Figure 39. Simulation of closed-loop transition detector with Vdd=5.5V and fast process parameters

trailing edge of the DATB pulse is more delayed for a higher power supply voltage. This is because LA or LAB starts out at a higher voltage and takes longer to pull down. This behavior is different from that of the open-loop detector. It is also different from the behavior of the decoders. The closed-loop detector must be adjusted to allow the decoders to precharge at low voltage, but at high voltage the detector wastes time by slowing down while the decoders speed up.

The second disadvantage is the extreme sensitivity of the closed-loop detector to processing variation. The width of the DATB pulse varies from 9nS for Vdd=4.5V and fast processing parameters to 100nS for Vdd=5.5V and slow processing parameters. This is in contrast to the width of the DAT pulse in the open-loop detector, which varies from 9nS to 36nS.

Figure 40 and Figure 41 illustrate that the closed-loop transition detector has a single-valued D.C. response. One might expect that the response would be different depending on the direction of the transition, because of the memory in the circuit. This is not the case because the transfer functions of A1 and A1B cross over above the threshold voltage (obviously, because A1B must get somewhat above a threshold to pull down A1.) In this region of the D.C. response both M60 and M62 will be on, holding LOAD high so that the latch will not exhibit memory.

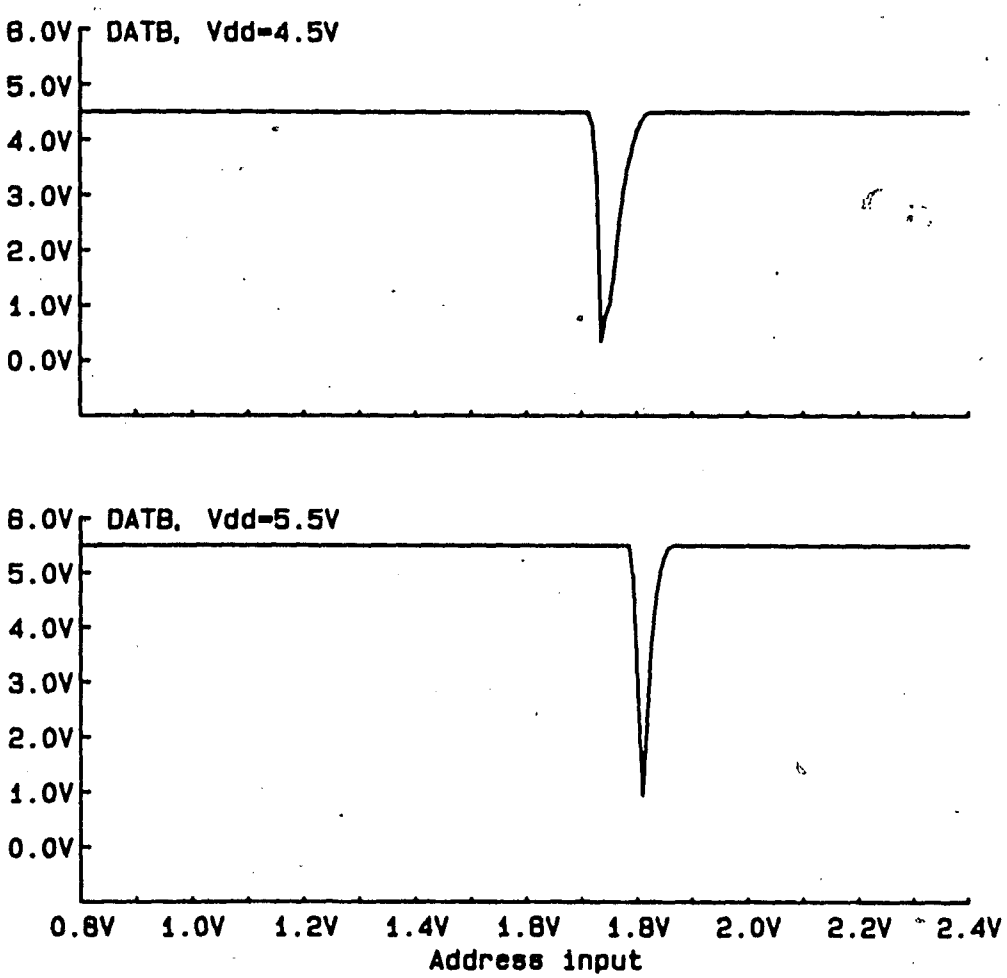


Figure 40. Simulated D.C. response of closed-loop transition detector with slow processing parameters

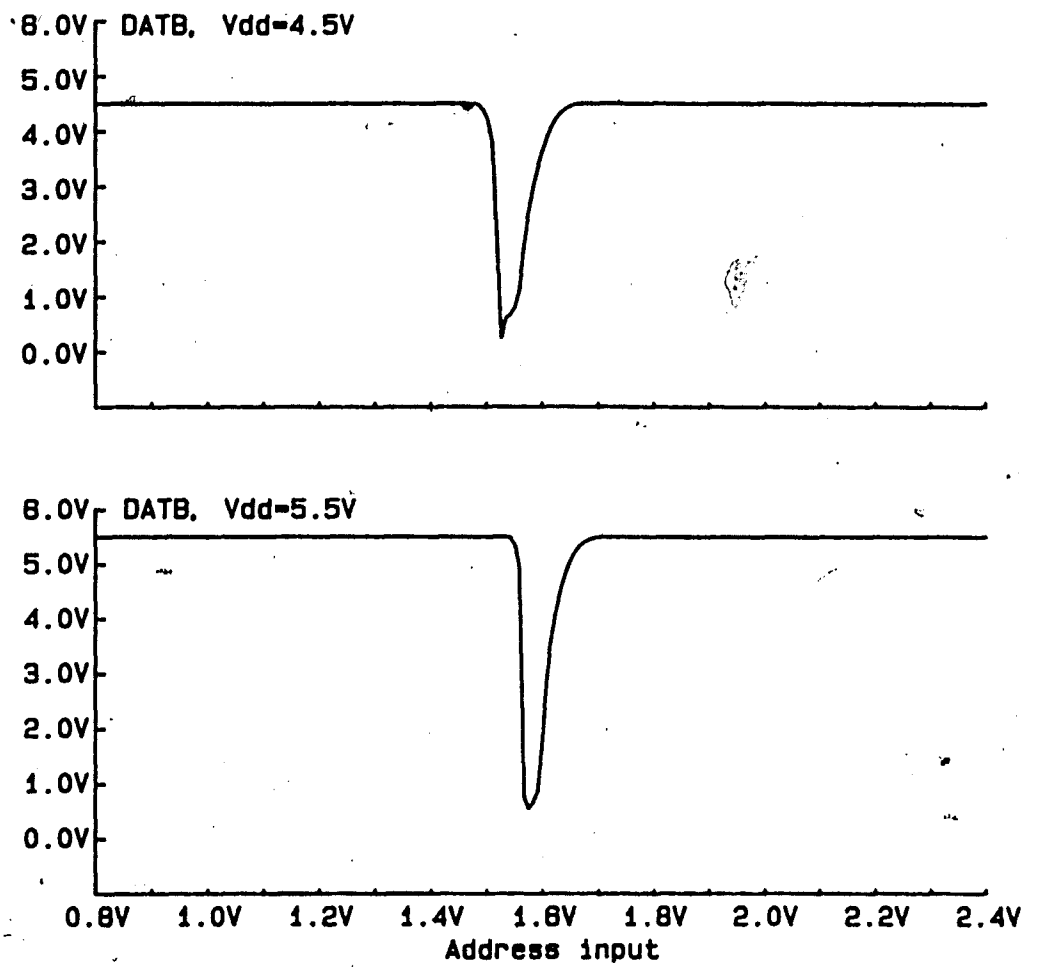


Figure 41. Simulated D.C. response of closed-loop transition detector with fast processing parameters

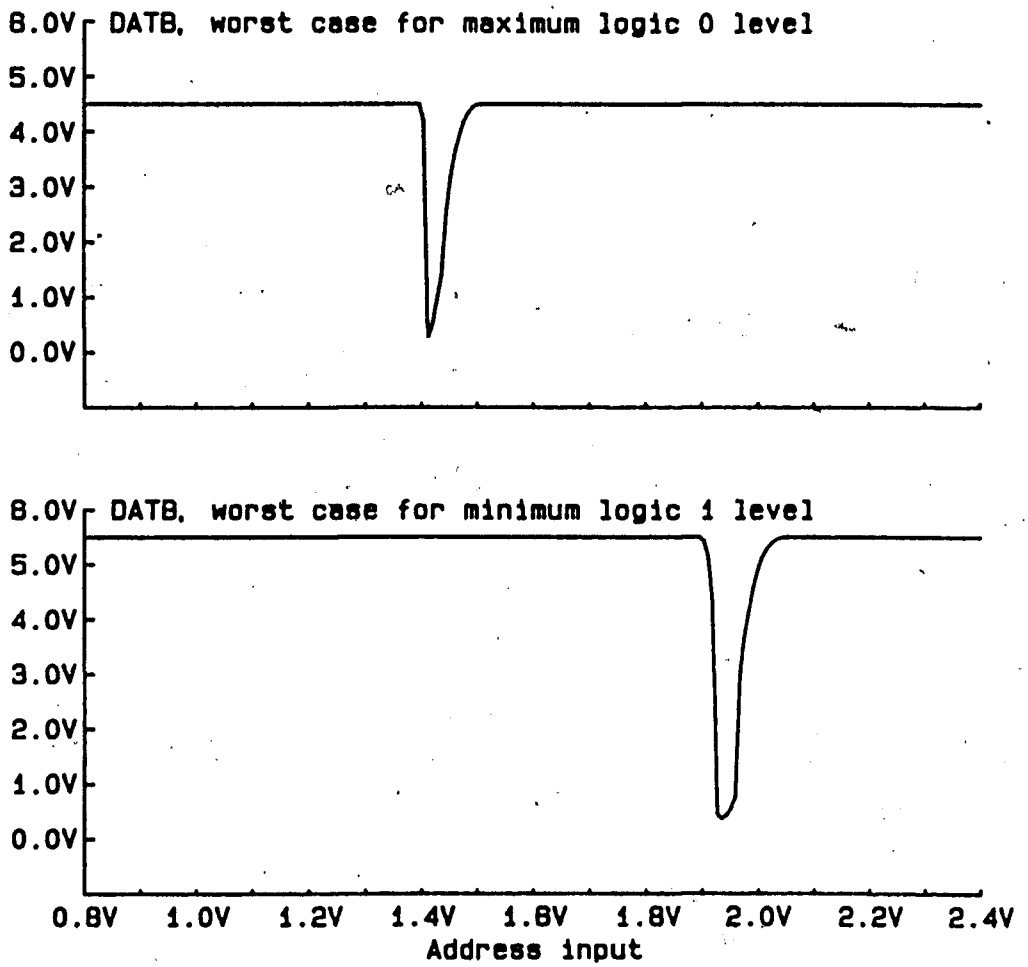


Figure 42. Simulated D.C. response of closed-loop transition detector for worst case input level margins

Figure 42 shows that under conditions to minimize the margins for the logic input levels, adequate safety margin exists.

Figure 43 shows the results of a Monte-Carlo analysis to determine the timing consistency of this detector. The definition of tAD is not changed from the open-loop circuit, except that the trailing edge of the DATB pulse is in a different direction than the trailing edge of the DAT pulse. The measurement is still made to the trailing edge. The graph shows that for aborted cycle operation, tAD changes by more than 20nS. The reason for this is that the latch is only partially loaded before the cycle aborts, and thus the latch changes back to its original state much more quickly. it should be noted that the implementation in (27), which derived the address lines from the output of the latch, would probably show much more consistency.

Summary. The open-loop transition detector operates properly and reliably for extreme ranges of processing and power supply levels, it responds to fast and slow transitions, and it has consistent timing. The circuit meets all of the design objectives. On the other hand, the closed-loop transition detector operates properly for extreme ranges of processing and power supply levels, and is provably reliable for any form of transition. However, it has extreme sensitivity to processing, reverse sensitivity to power supply level, and poor timing consistency. Fabrication of the circuits and experimental results will be given in the next chapter.

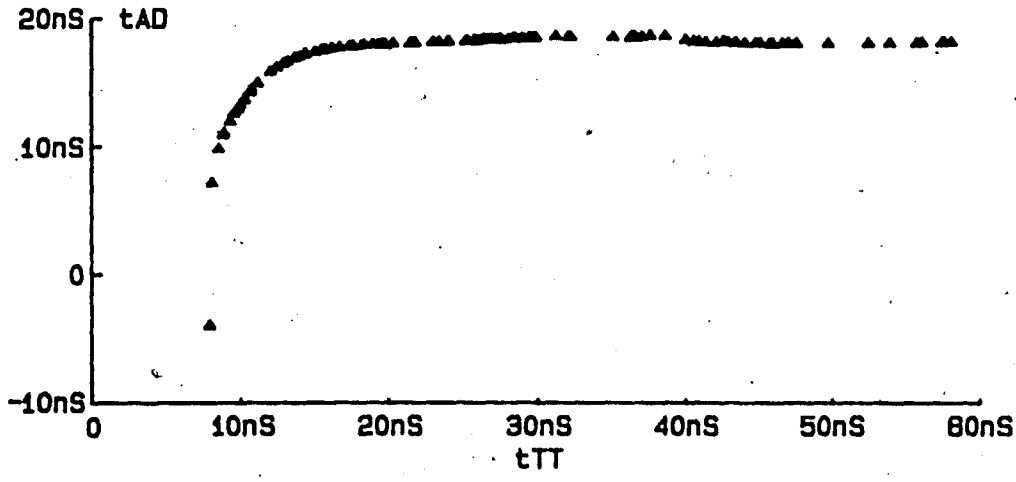


Figure 43. Simulation of closed-loop transition detector showing delay from address line to DATB (tAD) versus time between address input transitions (tTT) for Vdd=5.0V and nominal process parameters

4. Fabrication and Experimental Results.

Fabrication of each several samples of each these circuits was carried out by the Silicon Processing Technology Laboratory of Bell Telephone Laboratories, Inc. The open-loop detector was actually incorporated into a memory chip. The closed-loop detector was fabricated as a test circuit, with extra capacitance added to simulate the loading that would be present. This chapter reports the evaluation of these circuits on a randomly selected wafer for each implementation.

The actual processing parameters obtained on these wafers were measured on a Keithley LPT/2 automatic parametric measuring system. The actual processing parameters are shown in Table 3. These parameters were included in ADVICE simulations contained in this chapter.

Figure 44 and Figure 45 compare the characteristics of some test transistors (one for each threshold) on the closed-loop sample to the characteristics predicted by the model for the parameters of the closed loop sample. As noted in Chapter 3, the model contains inaccuracies for all of the devices. In particular, because it neglects sub-surface conduction, the model underpredicts the current in the depletion transistors for most conditions.

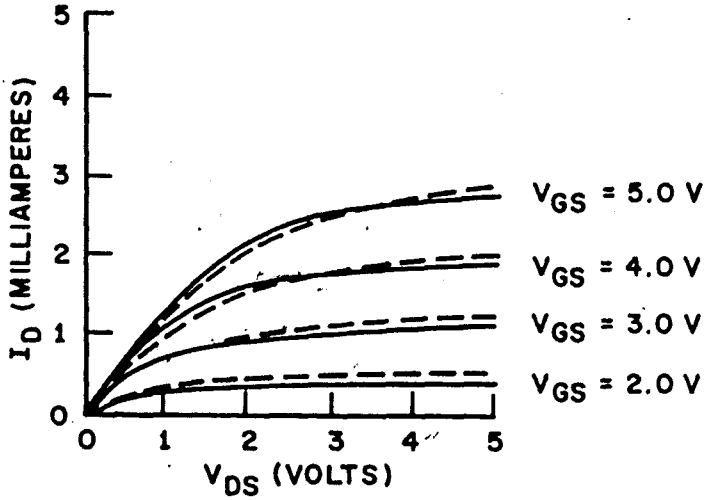
Measurements of circuit response were carried out using a Tektronix low-capacitance FET probe coupled by a small capacitor to a

Parameter	Nominal	Actual, Open Loop	Actual, Closed Loop
L′	2.0um	2.24um	1.70um
tOX	50nm	56.3nm	56.7nm
N	1.0×10^{15} cm^{-3}	8.62×10^{14} cm^{-3}	8.0×10^{14} cm^{-3}
VtH	1.15V	1.09V	0.93V
VtL	0.723V	0.58V	0.44V
VtN	-0.447V	-0.62V	-0.75V
VtD	-2.70V	-2.90V	-3.08V

Table 3. Actual Processing Parameters

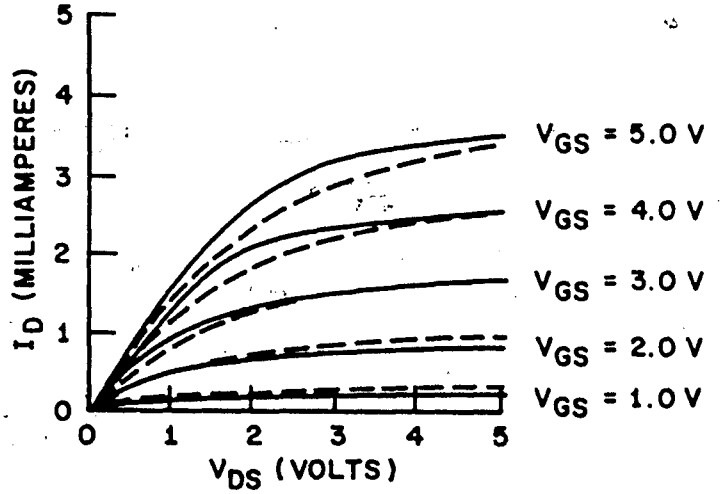
HIGH THRESHOLD ENHANCEMENT TRANSISTOR

W = 30μm L' = 1.70μm V_{BS} - 3.0V 28 °C



LOW THRESHOLD ENHANCEMENT TRANSISTOR

W = 30μm L' = 1.70μm V_{BS} - 3.0V 28 °C



——— MEASURED

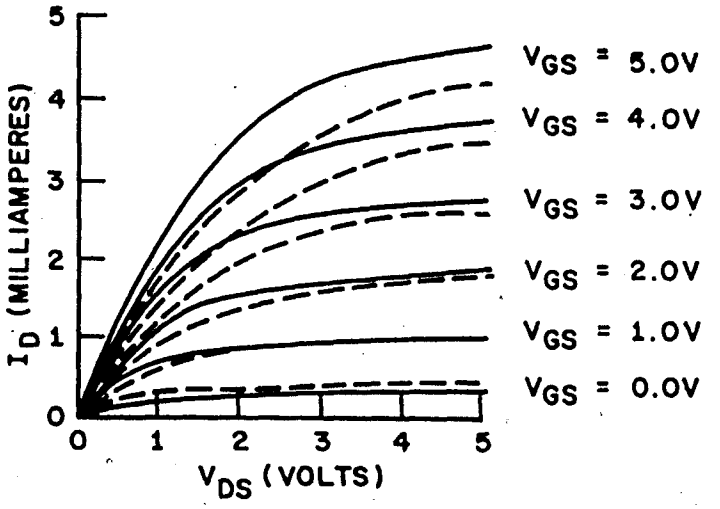
– – – – SIMULATED

COMPARISON OF MODEL WITH MEASURED I-V CURVES FOR ENHANCEMENT NMOS TRANSISTORS

FIGURE 44

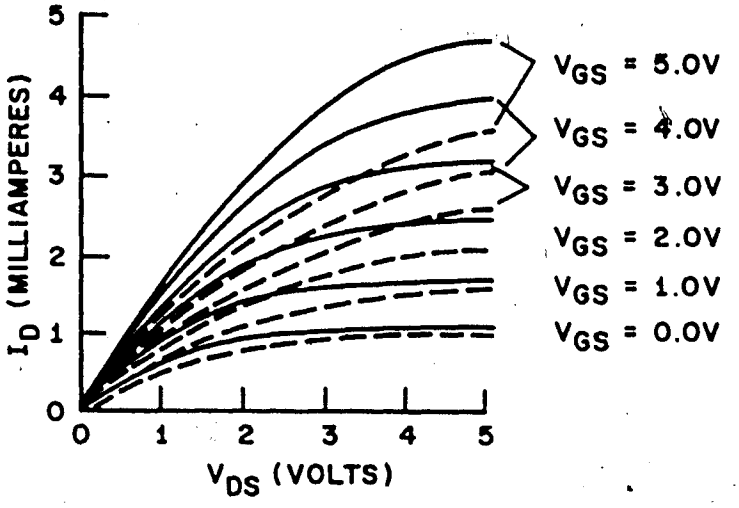
LIGHT DEPLETION TRANSISTOR

W = 30μm L = 1.70μm V_{BS} - 3.0V 28 °C



HEAVY DEPLETION TRANSISTOR

W = 30μm L' = 1.70μm V_{BS} - 3.0V 28 °C



——— MEASURED

– – – – SIMULATED

COMPARISON OF MODEL WITH MEASURED I-V CURVES FOR DEPLETION NMOS TRANSISTORS

FIGURE 45

tungsten probe wire of diameter less than 8 micrometers. The A.C. voltage gain of the probe was calibrated by comparison with another probe of known gain, using the signal on the address input as a reference. The effect of the probe on the internal timing was checked by comparing the response of the first clock with the probe on DAT to the response without the probe. The effect was found to be negligible.

The D.C. level of the signal shown in the measurements was assumed, since the FET probe was capacitively coupled. The waveforms were displayed on a Tektronix oscilloscope and photographed. The waveform was then digitized and converted to a form acceptable to ADVICE using the program DIGIT (28) and a Hewlett-Packard graphics plotter. After the comparison simulations were run, the digitized experimental waveforms were plotted with the simulated ones.

The Open-Loop Circuit. Figure 46 and Figure 47 show the measured response of the circuit to a 0-1 transition and a 1-0 transition, respectively, under nominal conditions. All measurements were taken at 90 degrees Centigrade. The illustrations also show a simulation of the circuit under identical conditions. As shown, the circuit performs basically as expected. There are discrepancies between the simulation and the measurement in the timing of the edges and the detailed shape of the waveform. These discrepancies are partially due to inaccuracies in the device modeling as explained in Chapter 3 and illustrated above.

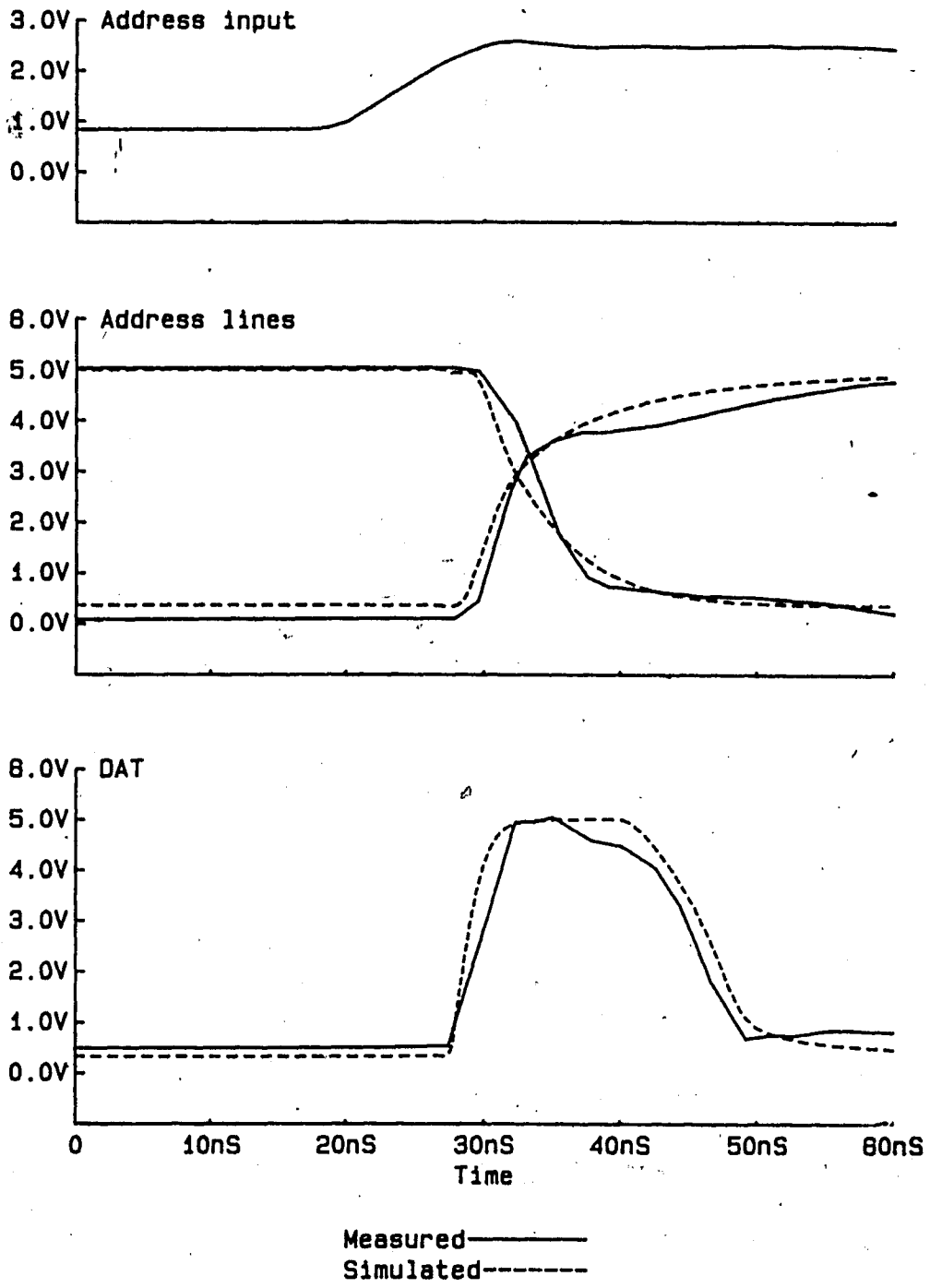


Figure 46. Operation of fabricated open-loop circuit compared with simulation at Vdd=5.0V for 0-1 transition

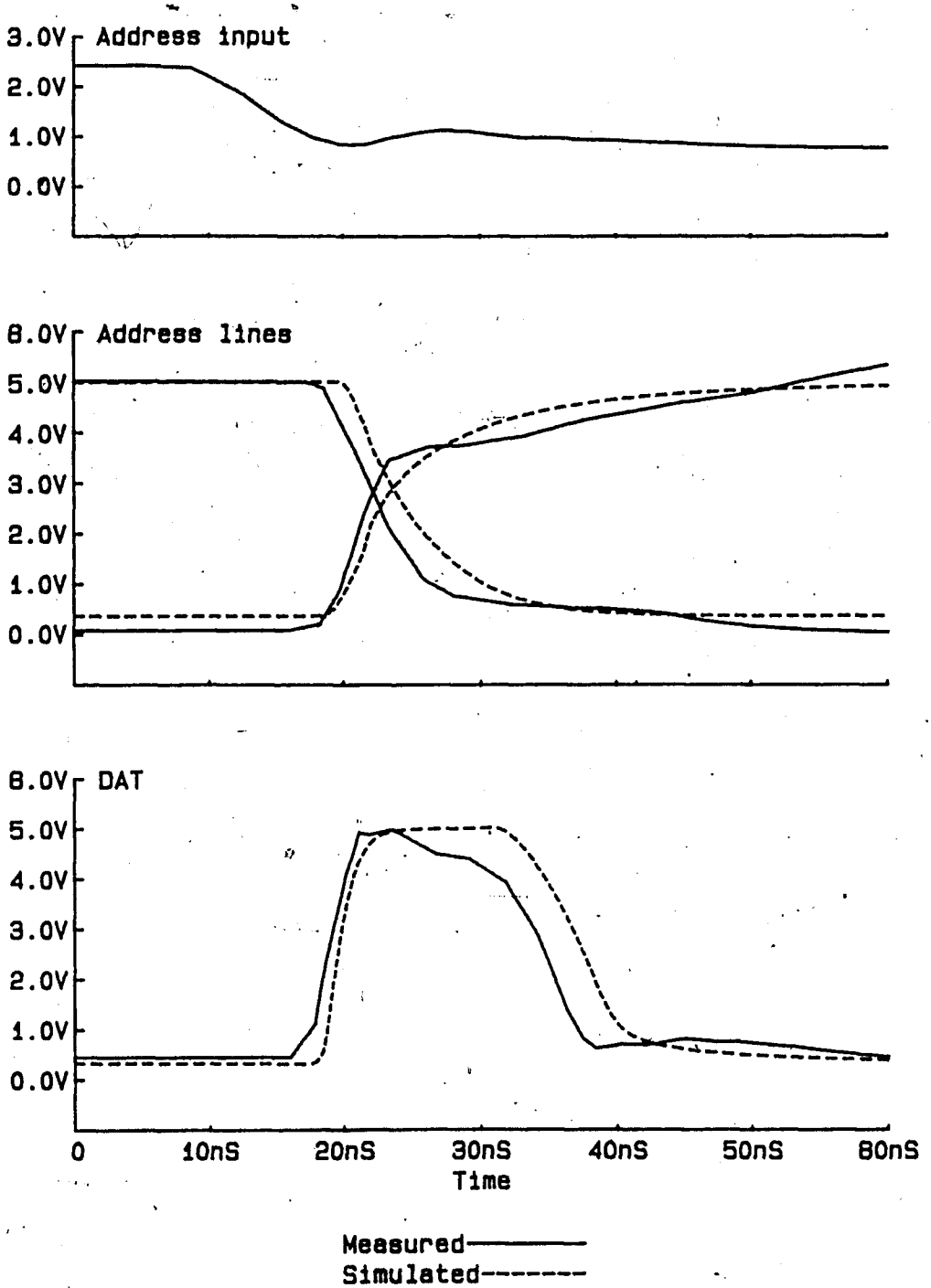


Figure 47. Operation of fabricated open-loop circuit compared with simulation at Vdd=5.0V for 1-0 transition

The discrepancies are also due to parasitic inductors in the integrated circuit and the measurement setup. These inductors, notably in the Vss power supply bus, are very difficult to measure and properly model. ADVICE does not contain the capability to model the frequency-dependent inductance of a metal stripe over a resistive ground plane, which the on-chip Vss bus is. Figure 48 shows the effect of adding some (roughly estimated) inductance to a nominal process simulation.

Figure 49 and Figure 50 compare the actual response to the simulation for extreme ranges of the power supply. A 0-1 transition only is shown. Again the circuit response is substantially as the simulation predicts.

Figure 51 shows the D.C. response of the circuit compared to a simulation. Here considerable discrepancy exists in which region of the input range the response occurs in. This is primarily due to the inaccuracy in modeling the depletion transistor. The first stage of the address buffer has a very high D.C. voltage gain (in order to amplify the TTL-level transition to full power supply levels) which magnifies the discrepancy in the pullup of this stage and causes it to offset the response. If the sub-surface conduction region did not exist, the pullup device would provide less current and the detector would respond at a lower level of drive on the address input, as predicted by the simulation.

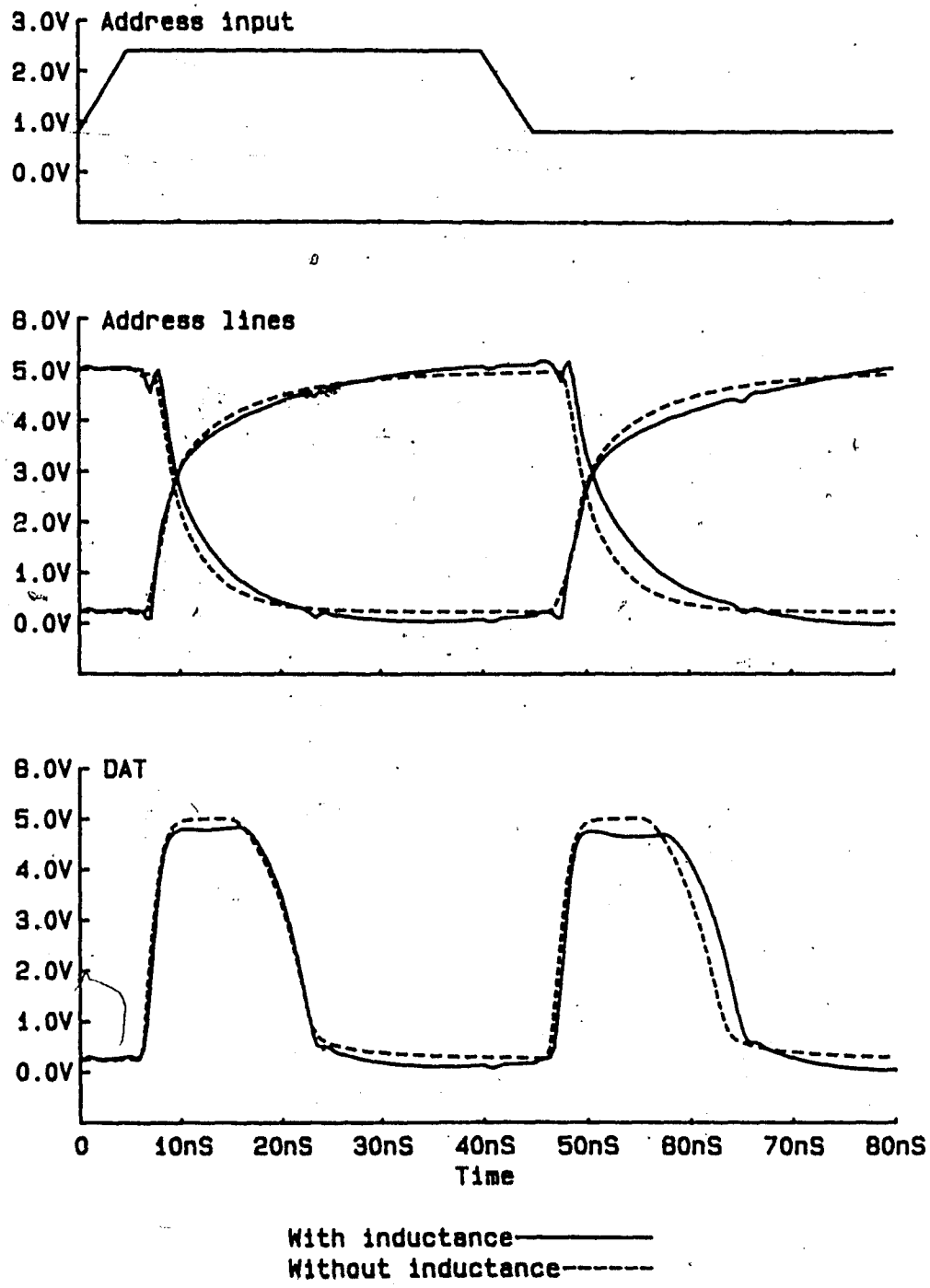


Figure 48. Effect of inductance on circuit simulation with nominal process parameters and Vdd=5.0V

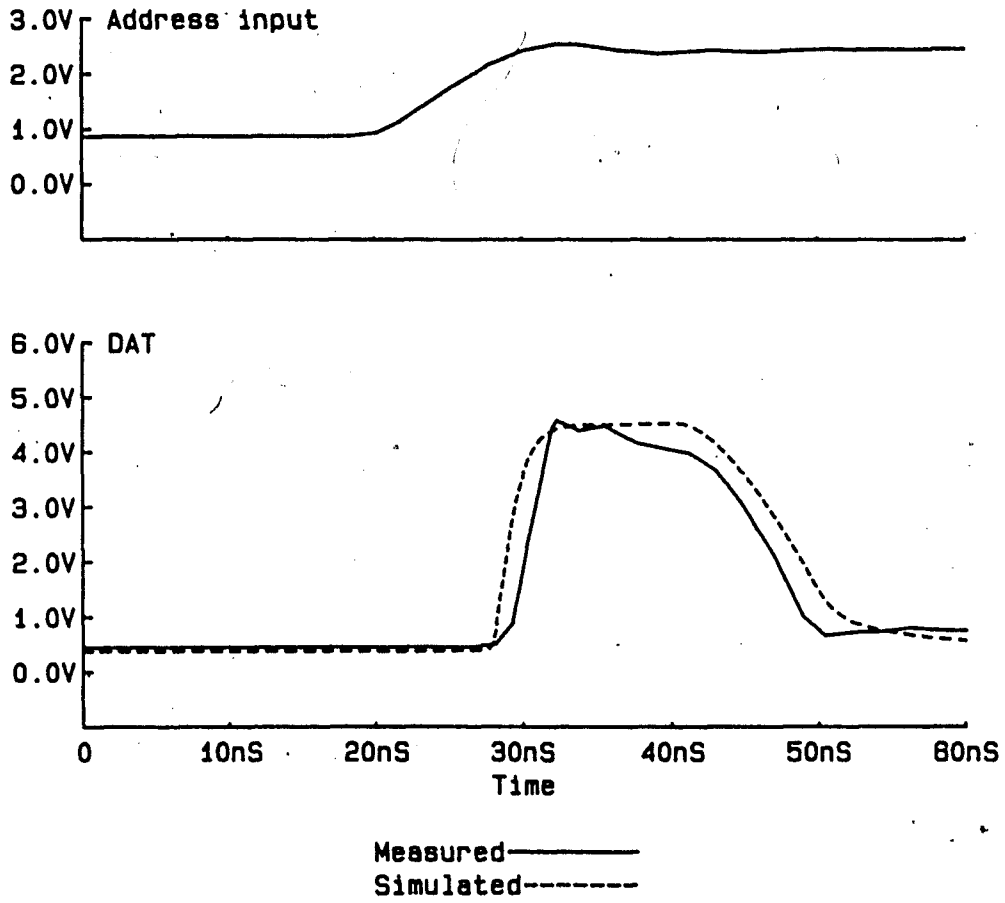


Figure 49. Operation of fabricated open-loop circuit compared with simulation at Vdd=4.5V

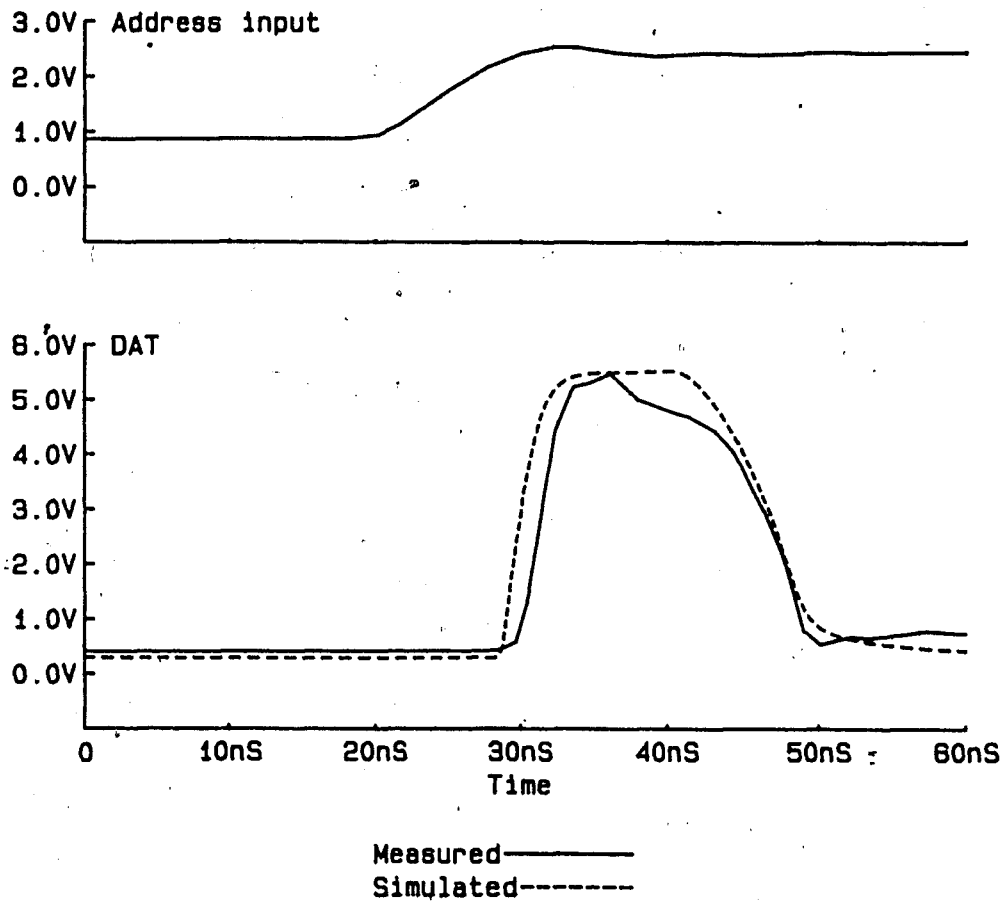
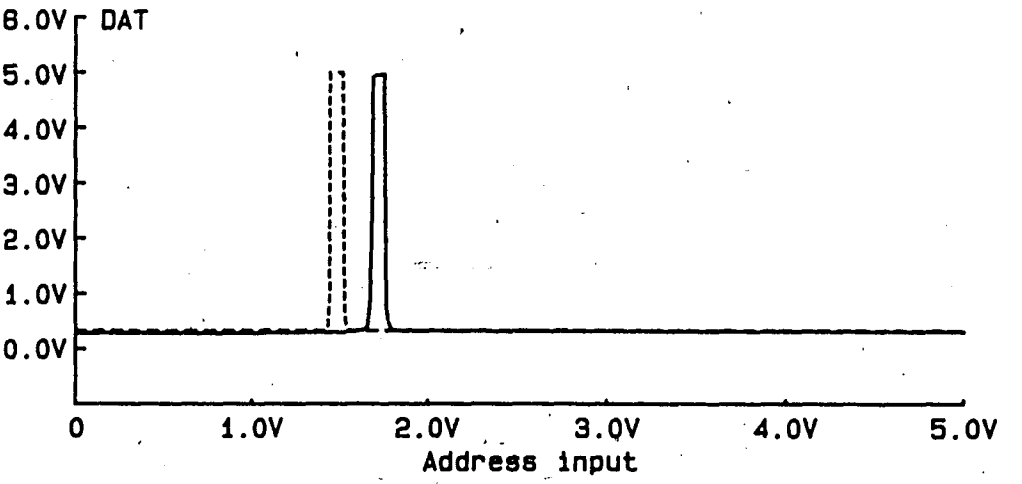


Figure 50. Operation of fabricated open-loop circuit compared with simulation at Vdd=5.5V



Measured————
Simulated-------

Figure 51. D.C. response of fabricated open-loop circuit compared with simulation for Vdd=5.0V

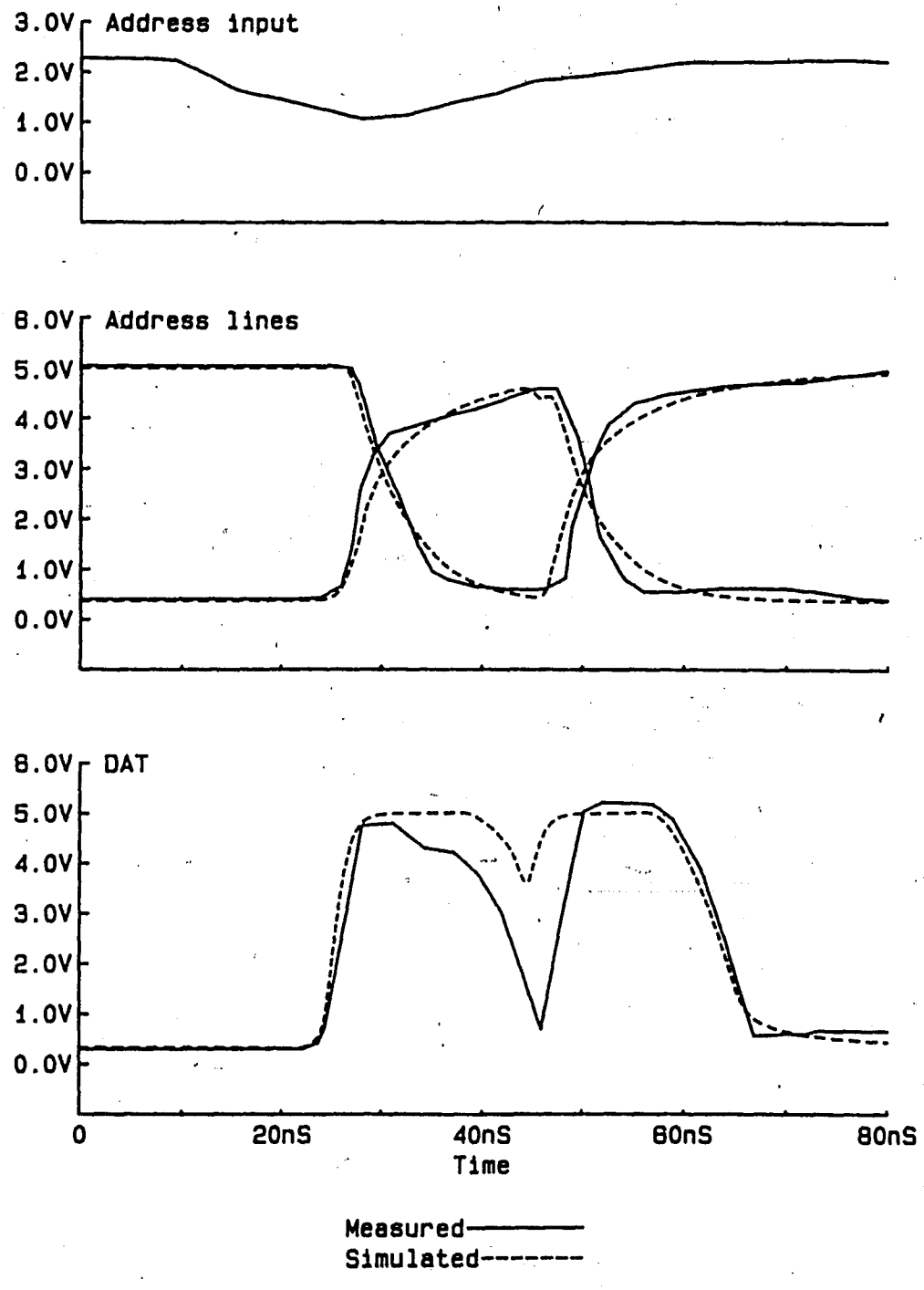


Figure 52. Operation of fabricated open-loop circuit compared with simulation at Vdd=5.0V for double transition

Figure 52 shows the response of the detector to a double transition. The simulation adequately predicts the time between the rising of the address line and the end of the DAT pulse.

The Closed-Loop Circuit. Figure 53 and Figure 54 show the measured response of the circuit to a 0-1 transition and a 1-0 transition, respectively, under nominal conditions. As before, the measurements were carried out at 90 degrees Centigrade, and a simulation under identical conditions is included. The most striking discrepancy here is the rate at which DATB recovers to Vdd. This discrepancy is probably due to the effect of the probe capacitance on DATB. The driving devices on DATB are actually smaller than the drivers of any of the nodes probed so far. Since this circuit is not actually part of a chip, it was impossible to measure the effect of the probe by measuring another clock output farther down the chain.

Figure 55 and Figure 56 show the response to a 0-1 transition, for extremes of power supply voltage. Again, except for the discrepancy already noted, the measurement agrees well with the simulation. The width of DATB is 19nS for Vdd=5.5V versus 15nS for Vdd=4.5V.

Figure 57 shows the D.C. response of the circuit. The same discrepancy exists as in the open-loop detector, again because of the poor depletion transistor model.

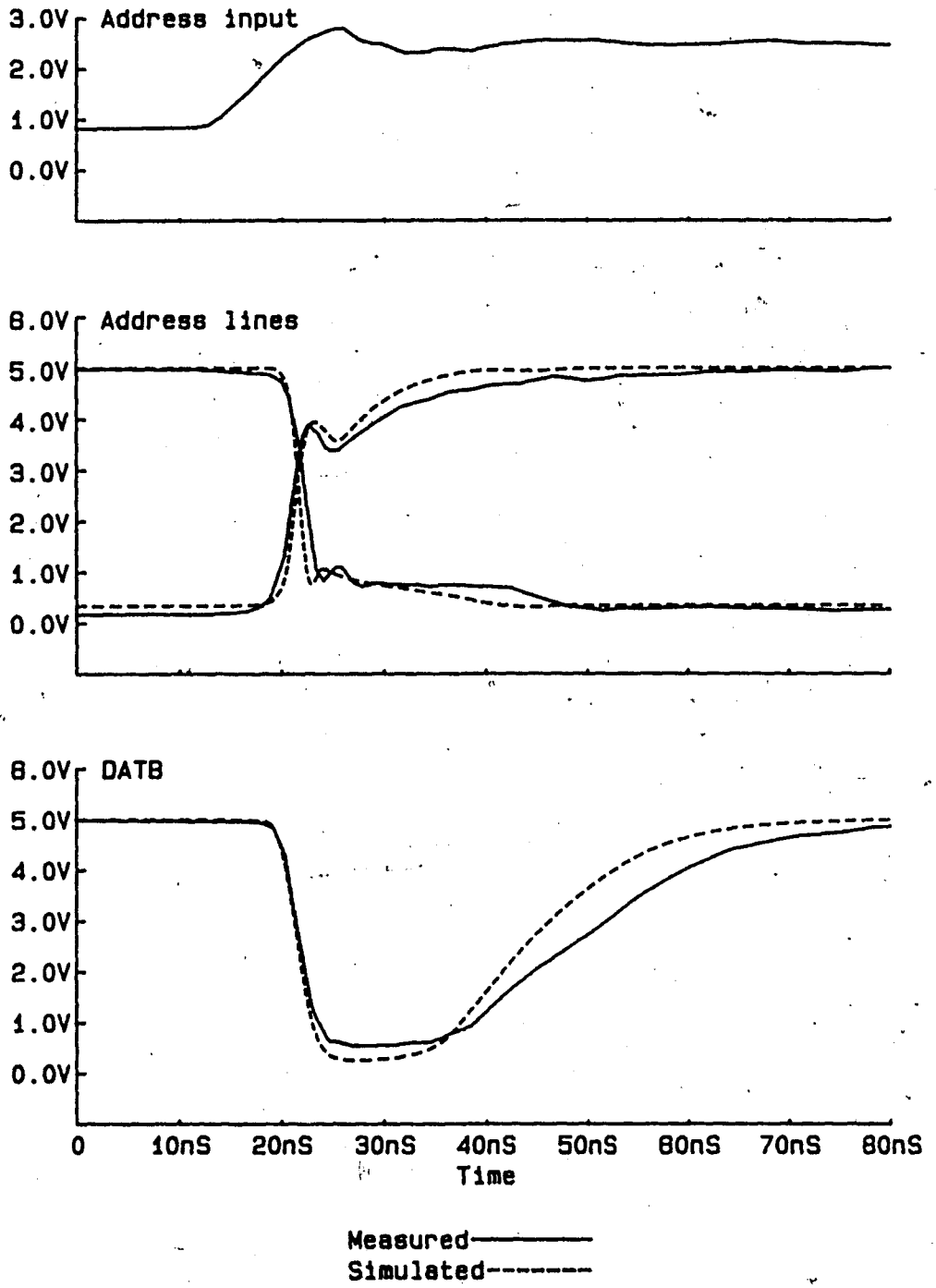


Figure 53. Operation of fabricated closed-loop circuit compared with simulation at Vdd=5.0V for 0-1 transition

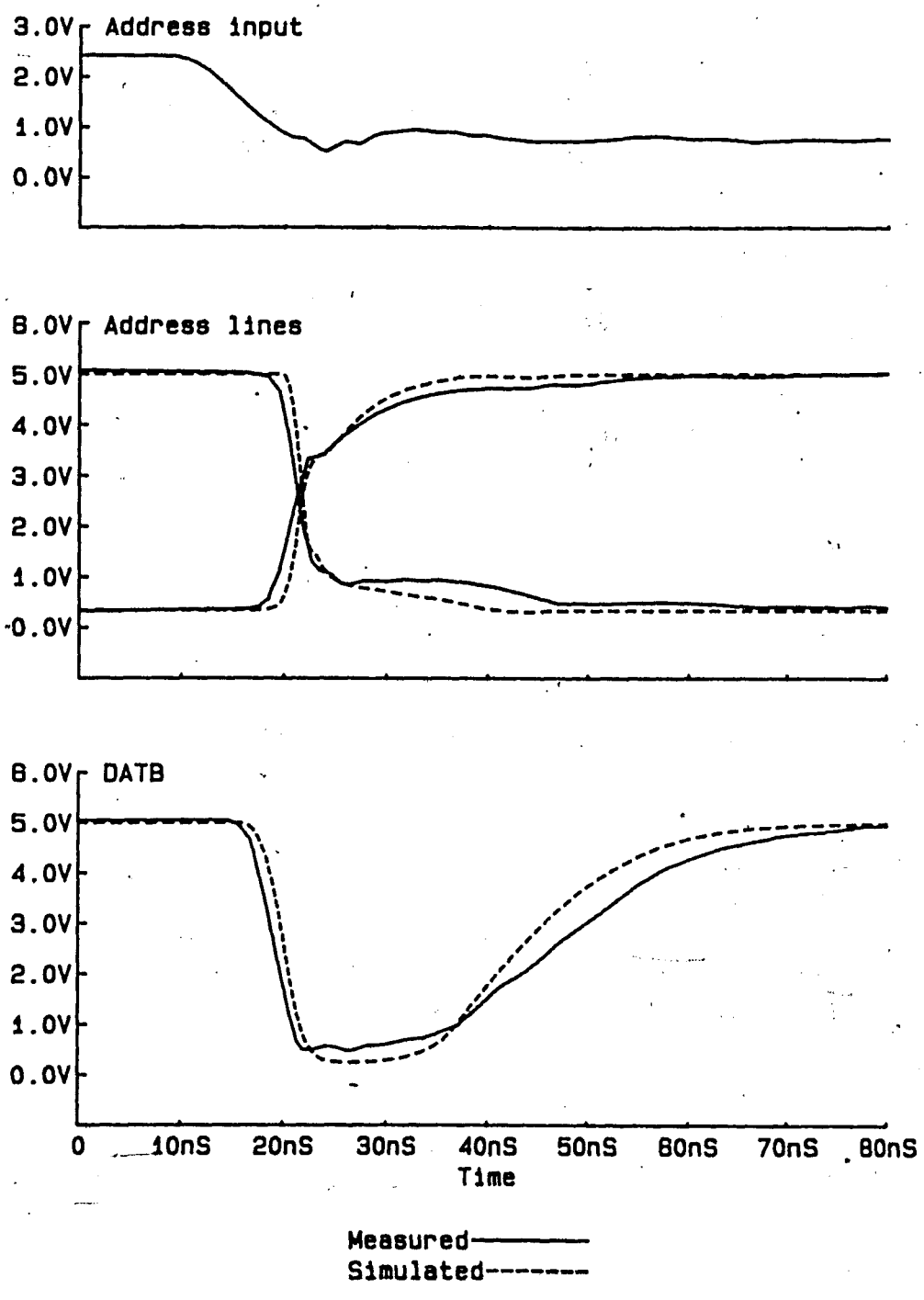


Figure 54. Operation of fabricated closed-loop circuit compared with simulation at Vdd=5.0V for 1-0 transition

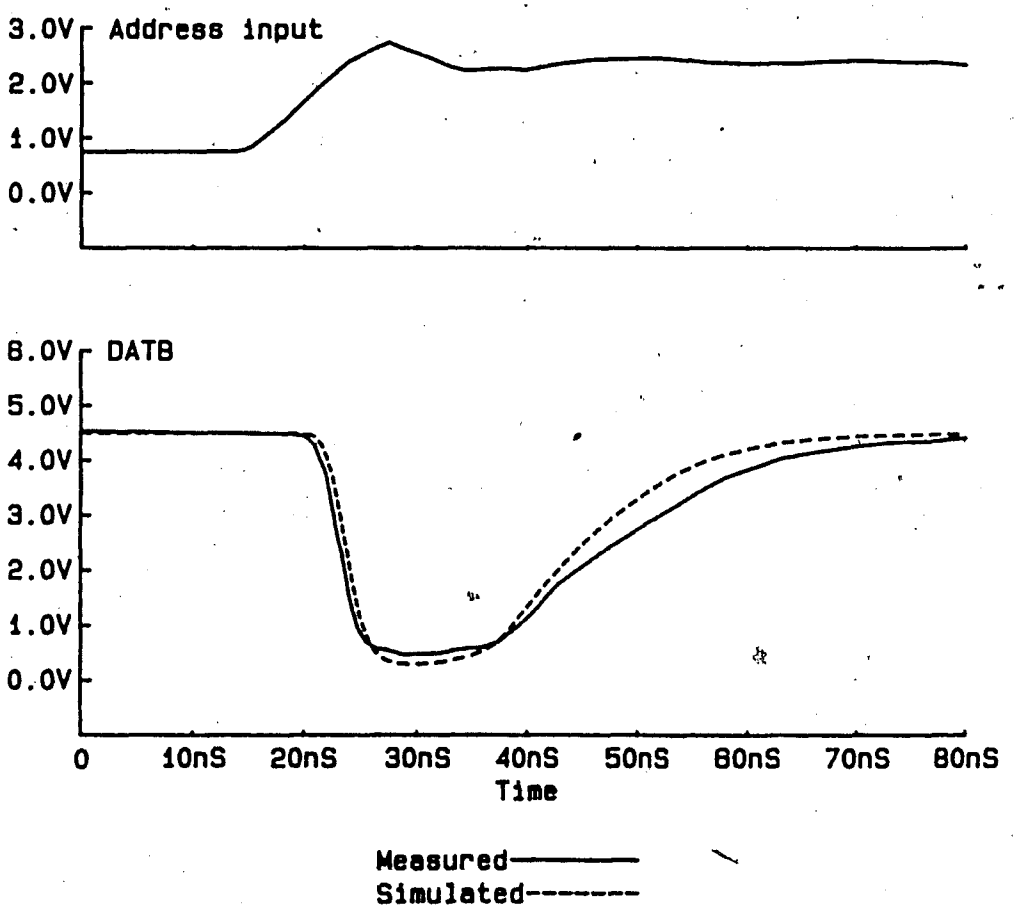


Figure 55. Operation of fabricated closed-loop circuit compared with simulation at Vdd=4.5V for 0-1 transition

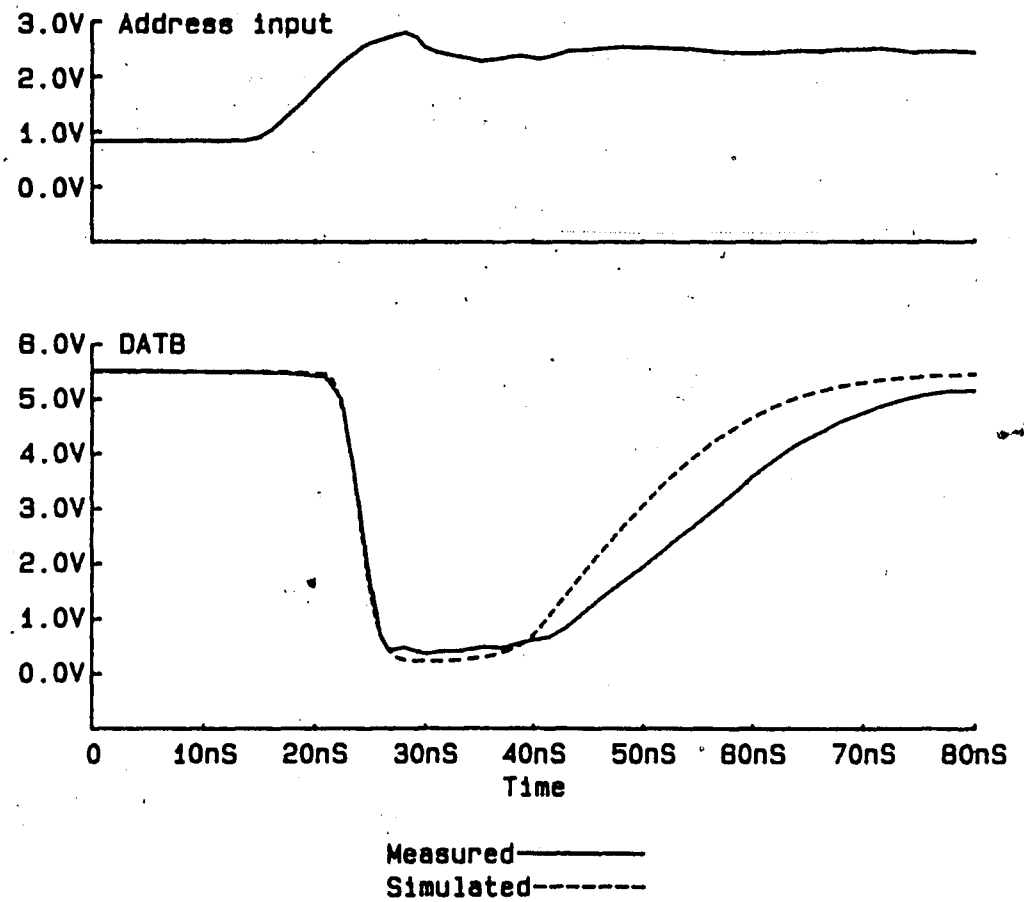


Figure 56. Operation of fabricated closed-loop circuit compared with simulation at Vdd=5.5V for 0-1 transition

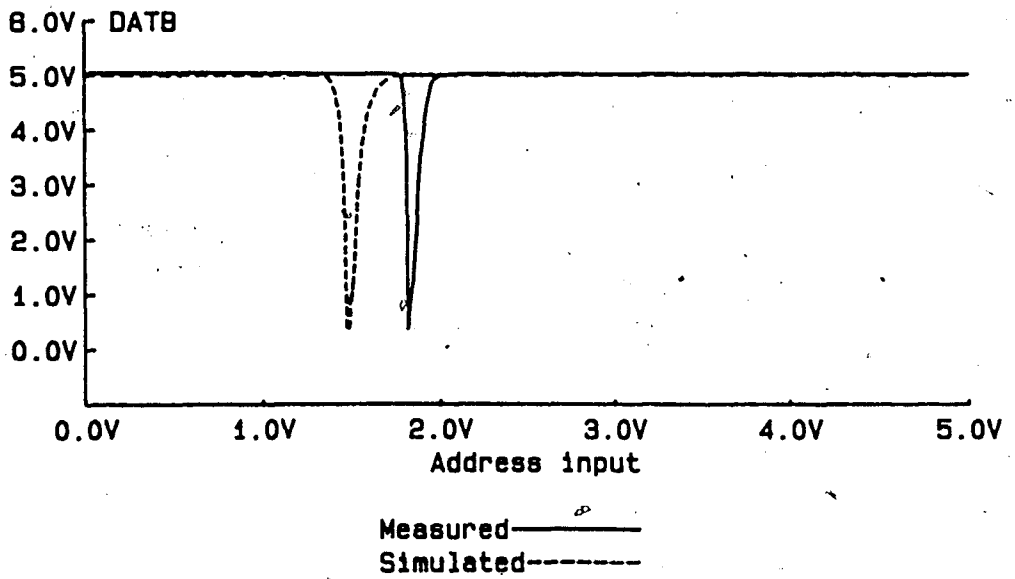


Figure 57. D.C. response of fabricated closed-loop circuit compared with simulation for Vdd=5.0V

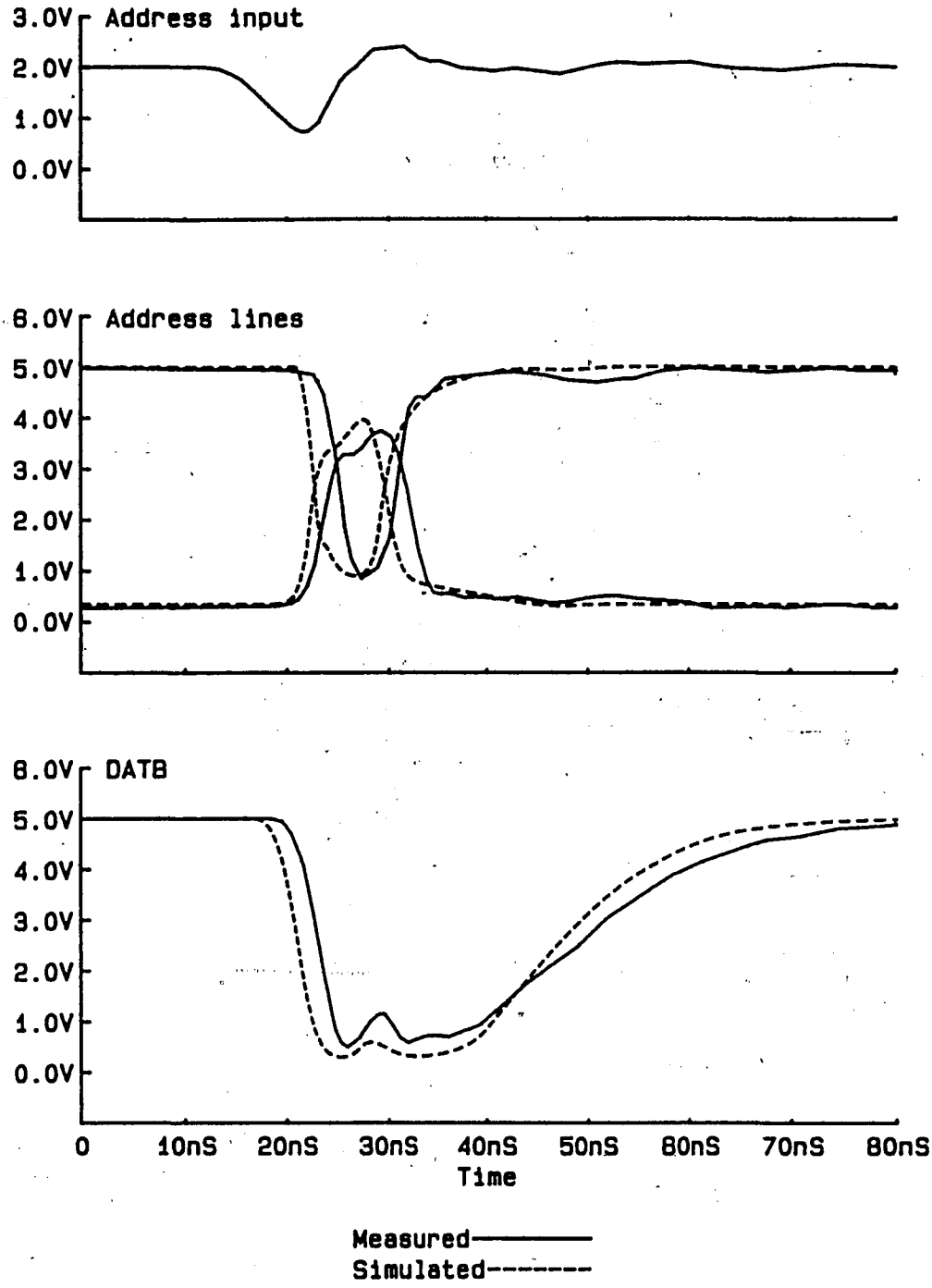


Figure 58. Operation of fabricated closed-loop circuit compared with simulation at Vdd=5.0V for double transition

Figure 58 shows the response to a double transition. Again good agreement with simulation is evident.

Summary. Fabrication of the open-loop transition detector and the closed-loop transition detector has been accomplished. In general the transition detectors behave as the simulations predict. Discrepancies exist due to the economies of modeling and simulation and the limitations of measurement. These discrepancies are not severe enough to influence the evaluation of the circuit.

5. Conclusion

Two designs for a transition detector have been evaluated and compared. The open-loop transition detector was developed from a simple and common edge detecting technique, resulting in a simple, reliable, and consistent circuit. The closed-loop transition detector was developed from a logical structure designed to guarantee an output pulse for any input transition. The resulting design was complex and performed inconsistently and poorly.

The Open-Loop Detector. The technique of comparing a delayed version of a signal with an undelayed version has been used in many circuits to detect edges or transitions. It is especially suitable for an integrated circuit application because of the tracking between different parts of the circuit of delay times. The NMOS implementation shows suitable performance for use in static memory chips. The circuit responds to an arbitrarily fast or slow transition and performs well for all extremes of processing and operational conditions. It also has excellent timing consistency for aborted cycle operation.

This open-loop transition detector has been incorporated into a static RAM chip which has been installed in many computer systems. It has performed reliably in these applications.

The Closed-Loop Detector. The design of the closed-loop transition detector derived from the desire to obtain a circuit which is provably reliable. The circuit stores the value of the input since the last transition and compares it to the input. A difference causes an output and the output is fed back into the circuit to store the new input.

The NMOS implementation of this circuit has design complications and several performance disadvantages. It is extremely sensitive to processing variation and shows a reverse sensitivity to power supply voltage. In addition, it has poor timing consistency for aborted cycle operation.

Recommendations for additional investigation. While the concept of the open-loop transition detector presented here should, with proper attention to details like D.C. response, yield a suitable, reliable circuit for most integrated circuit applications, much additional work is needed before our understanding of transition detector circuits can be complete. Continued alertness for conditions under which the open-loop detector may become unreliable is warranted. Other configurations of the closed-loop transition detector concept should be investigated for better performance and suitability. Novel concepts, such as those using differentiators, should be investigated.

List of References

1. Vadasz, Leslie L., Grove, Andrew S., Rowe, Thomas A., and Moore, Gordon E. "Silicon-gate Technology" I.E.E.E. Spectrum, Vol. 6 No. 10 (1969), p. 28.

2. _____. I.C. Master 1979. United Technical Publications, Inc. 1979.

3. O'Connell, Timothy R., Hartman, John M., Erret, Ellis B., and Leach, George S. "Two Static 4K Clocked and Nonclocked RAM Designs" I.E.E.E. Journal of Solid-State Circuits, Vol. SC-12 No. 5 (October 1977), p. 497.

4. Barber, F. E., Dumbri, A. C., Kane, J., and Rosenzweig, W., private communication, October 25, 1978.

5. Ohzone, T., Hirao, T., Tsuji, K., Horiuchi, S., and Takayanagi, S. "A 2Kx8 Bit Static RAM" Technical Digest 1978 International Electron Devices Meeting (1978), p. 360.

6. Jecmen, Robert M., Ebel, Anne A., Smith, Ronald J., Kynett, Virgil, Hui, Chi-Hung and Pashley, Richard D. "A 25nS 4K Static RAM" International Solid-State Circuits Conference Digest of Technical Papers, Vol. 22 (1979), p. 100.

7. Wada, Toshio, Kudoh, Osamu, Nagahashi, Yasuhiko, and Matsue, Shigeki "A 15-nS 1024-Bit Fully Static MOS RAM" I.E.E.E. Journal of Solid State Circuits, Vol. SC-13 No. 5 (October 1978), p. 635.

8. DePaolis, M. V., Dumbri, A. C., Hogg, R., Pantelakis, D. C., Serpiello, J. W., and Rosenzweig, W., private communication, May 1, 1981.

9. Jiang, Ching-Lin and Plachno, Robert "A 32K Static RAM Using a Three-Transistor Cell" International Solid-State Circuits Conference Digest of Technical Papers, Vol. 24 (1981), p. 86.

10. Iizuka, Tetsuya, Ochii, Kiyofumi, Ohtani, Takayuki, Kondo, Takeo, and Koyhama, Susumu "Fully Static 16Kb Bulk CMOS RAM" International Solid-State Circuits Conference Digest of Technical Papers, Vol. 23 (1980), p. 226.

11. Minato, Osamu, Masuhara, Toshiaki, Sasaki, Toshio, Nakamura, Hideaki, Sakai, Yoshio, Yasui, Tokumasa, and Uchibori, Kiyofumi

"2Kx8 Bit Hi-CMOS Static RAM's" I.E.E.E. Transactions on Electron Devices, Vol. ED-27 No. 8 (August 1980), p. 1591.

12. Masuhara, Toshiaki, Minato, Osamu, Sasaki, Toshio, Nakamura, Hideaki, Sakai, Yoshio, Yasui, Tokumasa, and Uchibori, Kiyofumi "2Kx8b HCMOS Static RAMs" International Solid-State Circuits Conference Digest of Technical Papers, Vol. 23 (1980), p. 224.

13. Stewart, Roger G. and Dingwall, Andrew G. F. "16K CMOS/SOS Asynchronous Static RAM" International Solid-State Circuits Conference Digest of Technical Papers, Vol. 22 (1979), p. 104.

14. Ohzone, Takashi, Kondo, Shuji, Tsuji, Kazuhiko, Shirigasawa, Tsuyoshi, Ishihara, Takeshi, and Horiuchi, Shiro "A 64Kb Static RAM" International Solid-State Circuits Conference Digest of Technical Papers, Vol. 23 (1980), p. 236.

15. _____. The TTL Data Book for Design Engineers. Texas Instruments, Inc., 1973.

16. DePaolis, Michael V. "A Strobed Data Output Buffer for a 2K X 8 Static NMOS Random Access Memory" Master's Thesis in Electrical Engineering, Lehigh University, 1980.

17. Clemens, J. T., Doklan, R. H., and Nolen, J. J. "An N-Channel Si-Gate Integrated Circuit Technology" Technical Digest 1975 International Electron Devices Meeting, 1975.

18. Nagel, L. W. "ADVICE for Circuit Simulation" I.E.E.E. International Symposium on Circuits and Systems, Houston, Texas, April 28, 1980.

19. Rosenzweig, W., Jones, J. R., and Serpiello, J. W., private communication, March 8, 1977.

20. Montgomery, Janet C. "A Comparison of Threshold Voltage Models for NMOS Transistors Including Short-Channel and Narrow-Width Effects" Master's Thesis in Electrical Engineering, Lehigh University, 1981.

21. El-Mansy, Youssef A. "Analysis and Characterization of the Depletion-Mode IGFET" I.E.E.E. Journal of Solid-State Circuits, Vol. SC-15 No. 3 (June 1980), p. 331.

22. Orton, Ralph R. "Single-gate transition detector responds to positive and negative edges" Electronic Design, Vol. 27 No. 14 (July 5, 1979), p. 84.

23. Liu, C. C., Mesples, P. M., and Millican, D. L. "Zero-Crossing Detector" I.B.M. Technical Disclosure Bulletin, Vol. 15 No. 5 (October 1972), p. 1432.

24. Bucek, F. J. and Malone, J. P. "Pulse (Glitch) Generator" I.B.M. Technical Disclosure Bulletin, Vol. 22 No. 8a (January 1980), p. 3185.

25. Huffman, David "Polysilicon-load RAMs plug into mainframes or microprocessors" Electronics, September 27, 1979, p. 131.

26. Bencivenga, M., private communication, October 3, 1978.

27. Stewart, Roger G. "High-Density CMOS ROM Arrays" I.E.E.E. Journal of Solid-State Circuits, Vol. SC-12 No. 5 (October 1977), p. 502.

28. Holder, Clinton H., Jr., private communication, July 9, 1981.

Vita

Mr. Richard Hogg was born in Scranton, Pennsylvania on April 29, 1958, the son of Mr. and Mrs. Bruce K. Hogg. He graduated from West Scranton High School in June, 1976. He received the Bachelor of Science Degree in Electrical Engineering in January, 1980 from Lehigh University, Bethlehem, Pennsylvania. He graduated with highest honors. Since January, 1980 he has been a Member of Technical Staff in the Memory Design Department of Bell Telephone Laboratories in Allentown, Pennsylvania. He resides in Allentown.